The Norwegian Housing Market
An Econometric Analysis with Regional Data

Bård Ola Tjønneland

Economic Theory and Econometrics

Department of Economics

University of Oslo

May 2014

University of Oslo

Master's Thesis

The Norwegian Housing Market

An Econometric Analysis with Regional Data

Author:
Bård Ola Tjønneland

Supervisor:
Ragnar Nymoen

Department of Economics

May 2014

The Norwegian Housing Market:
An Econometric Analysis with Regional Data

Bård Ola Tjønneland

http://www.duo.uio.no

Print: Reprosentralen, Universitetet i Oslo

. . . til Bestefar

Abstract

Norwegian housing prices have skyrocketed over the last two decades, with aggregate housing prices having soared upwards almost 400 percent since 1993. Based on an aggregate housing price model provided by Anundsen and Jansen (2013*b*), this thesis explores unchartered territory by incorporating regional data on housing prices and debt so as to capture and explain regional housing market developments. The data is aggregated into three regions, Oslo & Akershus, the South-West, and Northern Norway, over the period 1994Q1–2012Q4. This period allows for a study of how the regional markets fared in the face of the turmoil associated with the financial crisis of 2008.

The regional housing price model suggests that the housing markets in the various regions are remarkably synchronized with few regional differences. Due to the existence of an error correction term in the housing price relation, it seems that housing prices are in line with fundamentals. The results are supported by the robustness of the model and the significant test statistics. Moreover, the results suggest that the aggregate model does a good job in explaining regional trends in the housing markets. The thesis also manages to establish a strong linkage between housing prices and debt.

Preface

Caminante, no hay camino,
se hace camino al andar
Antonio Machado

This master thesis represents the final phase of my years as a student in economics. Apart from one semester at Toulouse School of Economics, I have spent the last five years at the Department of Economics, University of Oslo. I am grateful to all my fellow students for providing a stimulating and exciting environment throughout all these years.

The greatest acknowledgment is due to my supervisor, Ragnar Nymoen, Professor at the University of Oslo. His support, insightful comments and suggestions have been indispensable. Thanks also to André K. Anundsen at Norges Bank and Eilev S. Jansen at Statistics Norway for supplying me with the data, and for their attentiveness and helpful advice with regards to the replication study. Roger Bjørnstad and Samfunnsøkonomisk Analyse deserve thanks for providing me with the necessary data on regional housing prices. Fredrik Kostøl and Marcus Gjems Theie deserve special thanks for proofreading the thesis. Lastly, I wish to express my gratitude to family and friends for their support and encouragement. The thesis would not have been the same without the contribution of these people, but they are of course in no way responsible for any errors or inaccuracies.

Oslo, May 2014,

Bård Ola Tjønneland

Contents

List of Tables

List of Figures

1 Introduction

Most research on housing markets today focus on an aggregate level analysis. While important in itself to determine the general developments in housing markets, nuances are lost when regional dynamics are left out. This thesis seeks to investigate and examine the regional housing markets in Norway, and the determinants for what drive the housing prices on a regional level. It is beyond the scope of this thesis to conduct an analysis of all the regions in Norway, and therefore I have restricted myself to look at three regions: Oslo & Akershus, the South-West,[1] and Northern Norway.[2] It is the aspiration of this thesis that it can help explain regional housing market dynamics, and thus, lay the groundwork for further research on regional housing markets.

Due to data limitations it is currently difficult to conduct a full-fledged model of regional housing prices. Hence, in order to proceed with the analysis I will base my regional model on an aggregate model inspired by Anundsen and Jansen (2013*b*)[3] with accompanying (but updated) data. By using aggregately measured data when needed, the problem of not being able to collect and obtain all the relevant regional variables is sidestepped. The regional aspect in the model is captured by substituting e.g. housing prices and debt in the benchmark aggregate model with the corresponding (and adequately measured) variables on regional level. In this manner, I am able to infer how well the aggregate model succeeds in explaining regional housing prices, as well as – by including and substituting additional regional variables in the model – it can tell us how much certain regional variables explain housing prices in different regions.

The Norwegian housing market has been eagerly debated in recent years, with analysts reaching disparate conclusions. Even a couple of Nobel laureates in economics – most notably Robert J. Shiller[4] and Paul Krugman[5] – have joined the debate by sounding the alarm of a housing bubble in Norway. Bubble or not, it is a fact that housing prices in Norway almost have had a fivefold increase over the last twenty years, see Figure 1(a), while the average price level in the economy (as measured by the unadjusted CPI) barely has increased by one eight compared to this.[6] When seen in context with the housing price developments in other countries, see Figure 1(b), the growth in housing prices in Norway appear even more impressive. Besides the strong and continuous housing price

[1] Consisting of the two counties Rogaland and Hordaland.

[2] Nordland, Troms and Finnmark.

[3] This thesis is in large part based on the discussion paper which is an extended version of the published paper Anundsen and Jansen (2013*a*). Additionally, it featured in Anundsen (2014).

[4] Dagens Næringsliv, January 11, 2012, "Ekspert frykter norsk boligboble."

[5] Dagens Næringsliv, January 7, 2014, "Advarer mot norsk boligboble."

[6] Source: Statistics Norway, Table 07230 and Table 03013.

growth, the housing market has shown a remarkable resilience in the face of one of the most tumultuous economic periods in the post-war era, with a startling growth of over 40 percent since the trough in 2008.[7] A similar boom in the housing market has not been experienced in Norway since the financial markets were deregulated in the 1980s. On that occasion it ended with a crash of the housing market and a full-blown banking crisis, see e.g. Vale (2004).

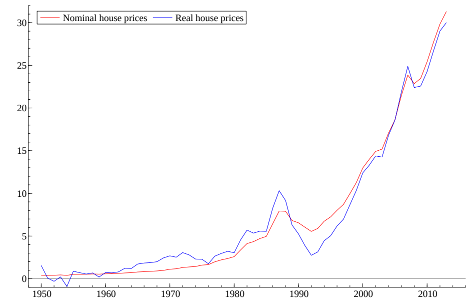
The housing market is of special importance to the economy due to the way it is intertwined with the banking sector, and the corresponding notion of financial stability. Financial intermediaries grant credit to individuals and households on the basis of debt-servicing capacity and the collateral posed, which means that when housing prices soar and households' net worth increases, credit becomes more accessible. This pro-cyclical pattern of credit availability can lead financial instability to build up in the economy, which makes it (and the households it comprises) more vulnerable to macroeconomic shocks. This is especially apparent in Norway, where mortgages constitute 90 percent of households' debt burden (Borgersen and Hungnes, 2009), and household debt stands at about 200 percent of disposable income (NOU, 2011; IMF, 2013), indicating that any (small) macroeconomic shocks that affect the housing market could amplify and have persistent effects on the real economy through a financial accelerator effect, as described in the seminal papers by Bernanke et al. (1999) and Kiyotaki and Moore (1997). This became painfully evident when American homeowners started defaulting on their mortgages in the mid 2000's, eventually culminating in the crash of the financial system and the worst economic downturn since the Great Depression. As Gerdrup (2003, pg. 30) notes, there exists "a strong causal link between financial fragility and banking crises." Anundsen and Jansen (2013*b*) investigate the relationship between housing prices and household debt in Norway, and find evidence of self-reinforcing effects between the two. In relation to this, Figures 1(c) and 1(d) depict evolution the price-to-income, debt-to-income, and debt-to-price ratios in Norway since 1980.

Others have argued that there are good reasons for why housing prices are high in Norway. An economy running at (or near) full capacity with low unemployment and interest rates, point in the direction of higher prices. The central question, however, is how much of the increase in housing prices that can be explained by so-called *fundamentals*, e.g. income, wealth, interest rates, and how much is due to psychological and speculative factors – coined irrational exuberance, cf. Shiller (2000). This is closely related to Stiglitz's definition of a bubble: "if the reason that the price is high today is *only* because investors believe that the selling price will be high tomorrow – when "fundamental" factors do not

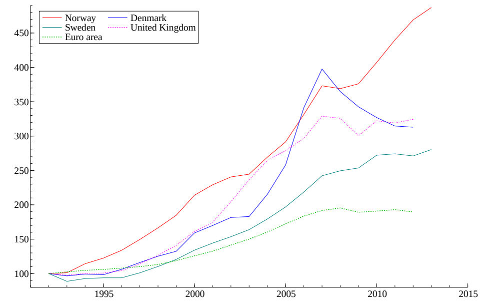
[7]Source: Statistics Norway

seem to justify such a price – then a bubble exists (Stiglitz, 1990, pg. 13).

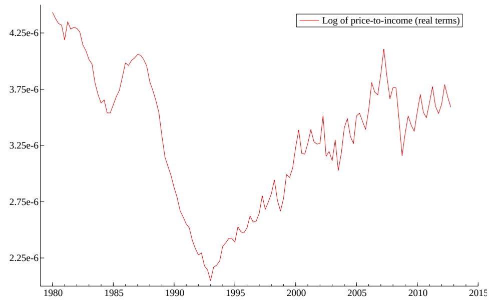
The thesis consists of eight chapters. Chapter 2 provides the theoretical framework laying the basis of the analysis, while the methodology and data description is presented in chapter 3. Central concepts of time-series econometrics are introduced in chapter 4, before a short literature review is offered in chapter 5. A replication and re-estimation of Anundsen and Jansen's (2013*b*) empirical study of the self-reinforcing effects between housing markets and credit is contained in chapter 6. A regional model for housing prices is presented in chapter 7, along with corresponding findings and results.[8] Chapter 8 concludes.



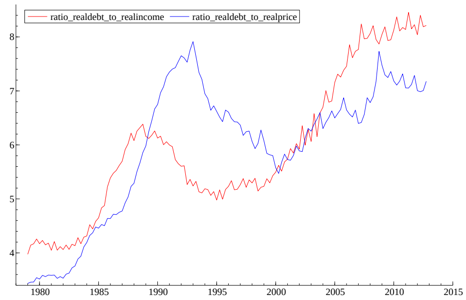
(a) Nominal and real housing prices



(b) Housing price indices (1992 = 100)



(c) Price-to-income (real terms)



(d) Debt-to-income and debt-to-prices (real terms)

Figure 1: The evolution of housing prices and debt. (*Source: Statistics Norway, Statistics Denmark, Statistics Sweden, Office for National Statistics, and Eurostat*)

[8] All the estimation- and test results are obtained by the statistical software OxMetrics7.

2 Theoretical Framework

The theoretical framework that underlies my analysis of which factors determine regional housing prices is inspired and based upon Anundsen and Jansen (2013*b*), which employ the housing model that is used in Statistics Norway's macroeconomic models MODAG and KVARTS. The exposition in this chapter follows a related work by Anundsen (2010) closely.

The recipe for defining a market is by no means simple (Tirole, 1988, pg. 12–13). The strand of research focusing on aggregate housing price models assumes one nationwide housing market for the whole country. It is easy to see that this is a strong assumption, as this *one* market consists of many smaller regional markets. However, it is a routine assumption to make, either because the purpose of the research project makes the assumption superfluous, or because it is the best we can do due to data limitations making it impossible to satisfy. One advantage by analyzing regional markets is that it allows the researcher to apply a richer and more specific information set to base the estimation, tests, and inference on that accounts for regional variation in the series.

2.1 Housing Demand

Housing demand, as emphasized by Jacobsen and Naug (2004*a*), consists of two components: households' demand for housing for consumption and investment purposes. In Norway, it is the former that accounts for the bulk of housing demand. Furthermore, there are two ways in which one can consume housing services: renting or owning. In the forthcoming analysis I will concentrate on the demand of housing services for owner-occupied dwellings.

A standard starting point in the literature is the following aggregate demand function for housing:

$$H^D = f(\underset{(-)}{PH}, \underset{(+)}{YH}, \underset{(+)}{D}, \underset{(-)}{R}, \mathbf{Z}) \tag{2.1.1}$$

where PH denotes the real housing price, YH is households' real disposable income, D is real debt, R is the real interest rate, and $\mathbf{Z}$ is a vector of other factors affecting the demand for housing services. The signs below the arguments in equation (2.1.1) signify the sign of the partial derivatives of a marginal increase in the respective argument *ceteris paribus*.

Under the assumption that housing services is an ordinary good, we expect that a higher price on housing services will lead to a decrease in demand. Furthermore, we expect that an increase in households' income will increase the demand for housing, thus treating housing as a normal good. As credit becomes more accessible, and household debt increases, demand is likely to increase as households can afford to pay more for housing services. An increase in the real interest rate is associated with a decrease in the demand for housing as the opportunity cost of housing increases. This can be justified in two different ways: (i) it becomes more expensive to borrow due to a higher interest burden, i.e. households must use a higher share of their income to pay down on their loans, thus, leaving them with less money to spend on other things, and (ii) it becomes relatively more profitable to deposit money in the bank instead of borrowing (which is particularly apparent for investors looking to make housing investments, as opposed to buying housing for consumption purposes, as a higher interest rate increases the required rate of return on the investment). Finally, the vector **Z** captures "everything else" that affect the demand for housing, e.g. demographic developments, costs of housing, and expectations about the future.[9]

Anundsen and Jansen (2013*b*) follow Jacobsen and Naug (2004*a*) in defining the user cost of housing as the value of a composite consumption good the household must forgo in order to own (and consume) one unit of housing for one period, but they augment the operational definition by adding a term capturing the presence of credit constraints. This is e.g. consistent with Meen and Andrew (1998) who rationalize the inclusion of such a term on the basis of banks' lending practices: The amount of credit made available by banks depends on debtors' net-worth. For instance, one type of credit constraint is present if households face an income constraint, namely that lenders set a maximum loan as a multiple of current income. Thus, the real user cost of housing is written as

$$HC_t = (1-\tau_t)i_t - \pi_t + \delta_t - \frac{\dot{PH}_e^t}{PH_t} + \lambda_t/\mu_c \tag{2.1.2}$$

where $R_t = (1-\tau_t)i_t - \pi_t$ is the real after-tax interest rate, δ_t is the depreciation rate (or the rate of maintenance costs including property taxation), and $\frac{\dot{PH}_e^t}{PH_t}$ is the expected real rate of appreciation for housing prices. The last term in equation (2.1.2) captures the credit constraint, where the shadow price of the credit constraint, λ_t, is divided by the marginal utility of consumption, μ_c.

This life-cycle framework implies that we must have equality between the marginal rate

[9] This includes not only expectations about prospective housing prices, but also expectations about future income, costs of housing, interest rate etc.

of substitution (MRS) between housing and a composite consumption good, i.e.

$$\frac{U_H}{U_C} = PH_t\left[(1-\tau_t)i_t - \pi_t + \delta_t - \frac{\dot{PH}_e^t}{PH_t} + \lambda_t/\mu_c\right] = PH_t HC_t \tag{2.1.3}$$

Market efficiency is obtained when the following no-arbitrage relationship holds,

$$PH_t = \frac{Q_t}{(1-\tau_t)i_t - \pi_t + \delta_t - \frac{\dot{PH}_t^e}{PH_t} + \lambda_t/\mu_c} = \frac{Q_t}{HC_t} \tag{2.1.4}$$

where Q_t is the real imputed rental price for housing services. The market is efficient if the user cost associated with a given dwelling is equal to what it would have cost to rent a dwelling of similar quality (Anundsen, 2013). Anundsen and Jansen (2013*b*) follow Meen (2002) and Poterba (1984) and interpret equation (2.1.4) as an inverted demand function. Furthermore, Anundsen and Jansen (2013*b*) assume a constant depreciation rate,[10] and that Q_t, which is unobservable, is a function of households' real disposable income (excluding dividends), YH_t, and the stock of dwellings, H_t. The inverted demand function can then be written as

$$PH_t = f^{-1}\left(H^D, YH_t, R_t, \frac{\dot{PH}_e^t}{PH_t}, \lambda_t/\mu_c\right) \tag{2.1.5}$$

Thus, with a constant depreciation rate, the real user cost of housing is then split in two components: The real direct user cost measured by R_t (which is used as the operational measure for HC_t in the forthcoming analysis), and the expected real housing price appreciation (captured by including lagged real housing prices in the model). The latter component of the user cost is consistent with Abraham and Hendershott (1996) which argue that lagged housing price appreciation do not have permanent effects, but act as a "bubble builder" by magnifying housing price increases as they pick up momentum. Finally, household loans is used as a proxy for the unobservable λ_t/μ_c.

2.2 Housing Supply

As already noted, the supply of housing can be assumed fixed in the short-run.[11] OBOS, a major operator initiating residential construction in Norway, operates with timeframes

[10]This is consistent with the Norwegian National Accounts, where a constant depreciation rate is used for housing.

[11]This assumption of a perfectly inelastic supply is obviously an approximation. A more correct statement would be to say that the supply of housing is (very) inelastic in the short-run. However, it is a routine assumption to make.

spanning from 10-15 years on their housing projects (Larsen and Sommervoll, 2004), which highlights the fact that it takes time before a project is initiated until it is finalized. Additionally, the construction of new residential property amounts to one percent of the housing stock each year (NOU, 2002), thus, making any year-by-year change in the housing stock negligible.

In the KVARTS framework that underlie Anundsen and Jansen (2013*b*), housing starts is a function of housing prices, construction costs, and the cost of land (Boug and Dyvi, 2008).

$$HS_t = g(\underset{(+)}{PH_t}, \underset{(-)}{CC_t}, \underset{(-)}{LC_t}) \tag{2.2.1}$$

where HS_t denotes housing starts, PH_t is housing prices which are likely to increase the profitability of new construction projects, CC_t and LC_t are construction and land costs respectively, which both are assumed to reduce the construction of new residential property. The total supply of housing, i.e. the stock of dwellings, is then determined by the following relation:

$$H_t^S = (1 - \delta_t)H_{t-1}^S + HS_t \tag{2.2.2}$$

where H_{t-1}^S is the housing stock last period, and δ_t is the depreciation rate on the housing stock. Hence, the long-run supply of housing is found by combining equation (2.2.1) and (2.2.3),

$$H_t^S = h(\underset{(+)}{PH_t}, \underset{(-)}{CC_t}, \underset{(-)}{LC_t}) \tag{2.2.3}$$

2.3 The Housing Market in the Short and Long Run

Like any other market, equilibrium is found where supply equals demand.[12] However, while the housing supply is variable in the long- run, it is assumed fixed in the short-run, i.e. it is perfectly inelastic as depicted in Figure 2. Therefore, in the short-run, the housing market price must adjust to bring the demand for housing in line with the existing supply (see Sørensen and Whitta-Jacobsen, 2010, chap. 14.4). Thus, any changes in the factors determining the demand for housing in equation (2.1.1), will lead to a shift in the demand curve in Figure 2, and a new equilibrium will arise.

[12]For a synopsis of the determinants of housing demand and supply see Pirounakis (2013, pg. 212).

Over time, however, the housing stock will adjust to reflect the demand of housing, indicating that the supply of housing is endogenous in the long-run. The long-run housing market equilibrium is found by combining the inverted housing demand equation (2.1.5) with the housing supply equation (2.2.3), as illustrated in Figure 3.

Another fruitful approach would be to calculate the reduced form equation of housing prices, that is housing prices as a function of exogenous variables, by inserting equation (2.2.3) into the housing demand equation (2.1.5) by using the fact that supply equals demand in equilibrium, $H_t = H_t^D = H_t^S$:

$$PH_t = f^{-1}\big(h(PH_t, CC_t, LC_t), YH_t, R_t, D_t\big) \tag{2.3.1}$$

This thesis follows Anundsen and Jansen (2013*b*), and treats the supply of housing as exogenous. Thus, the inverted demand function approach used here, applies a modified version of equation (2.1.5),

$$PH_t = f(\underset{(-)}{H_t}, \underset{(+)}{YH_t}, \underset{(-)}{R_t}, \underset{(+)}{D_t}) \tag{2.3.2}$$

In the analysis that follows, a semi-logarithmic transformation of equation (2.3.2) is used, where lower case letters denote log-scale.

$$ph_t = \beta_{1,1}h_t + \beta_{1,2}yh_t + \beta_{1,3}R_t + \beta_{1,4}d_t \tag{2.3.3}$$

In the representation above, it has implicitly been assumed a given state of expectations regarding variables such as housing prices, household income etc. Yet, it seems reasonable that expectations should be endogenously determined in the model. For instance, expectations of large future capital gains on housing may lead to an immediate boost in current property prices (Sørensen and Whitta-Jacobsen, 2010). The formation of such expectations may in turn trigger speculative behavior as households bid up the price of residential property just because they think prices will be high tomorrow. Such a "bigger fool" investment strategy may, nevertheless, be rational, as long as one is able to sell to a greater fool before prices slump, and proves that housing bubbles in many regards can be likened to equity price bubbles.[13] Endogenously determined expectations may, thus, be one explanation for why housing markets tend to go through long cycles of boom and bust.

[13] Reinhart and Rogoff (2009) define both equity and real estate bubbles as asset price bubbles.

2.4 The Debt Equation

As mentioned in section 2.1, household loans, i.e. debt, is used as a proxy for capital restraints in equation (2.1.2). Anundsen and Jansen (2013*b*) define household debt as a function of the housing stock, housing prices, the interest rate, disposable income, and the housing turnover (TH_t);

$$D_t = v(\underset{(+)}{H_t}, \underset{(+)}{YH_t}, \underset{(-)}{R_t}, \underset{(+)}{PH_t}, \underset{(+)}{TH_t}) \tag{2.4.1}$$

The interest rate effect on household debt is straightforward: an increase in the interest rate would increase the cost of debt servicing, and lead to a decrease in household debt. The product of the housing stock and housing prices can be interpreted as the market value of housing, and so an increase in the market value will increase households' net worth (and the corresponding collateral they can put up). This increase in net worth make households appear more financially robust, hence, leading to an increase in credit that banks are willing to supply (Anundsen, 2010, pg. 12). The same argument holds for real disposable income: an increase in income increases households' net worth and debt servicing ability, thus, increasing debt.[14] As discussed in Jacobsen and Naug (2004*b*), the effect on debt of an increase in housing turnover depends on what type of housing turnover one looks at. It is common to distinguish between the purchase of new dwellings, first-time and last-time purchase of an existing dwellings, and turnover of existing dwellings between households which neither enter nor exit the housing market. For the former, and under the assumption that the buyer borrows money to pay for the dwelling,[15] it is reasonable to believe that the debt will increase. This is because the seller in this case usually is not another household that can repay on an existing loan. In the second case, there is allowed for entry and exit in the market for existing dwellings. The exiting party gets freed up funds, and if these are not used to repay debt, then the total debt level in the economy increases. This is also reasonable on the grounds that the exiting party may enter the market for new dwellings,[16] or it may be less debt-burdened relatively to the

[14] A remark is here appropriate with regard to the life-cycle hypothesis which is the commonly used framework for modelling housing prices. If the life-cycle/permanent income hypothesis were true, one might actually observe an increase in the debt level as income falls. This is due to the fact that households want stability in their consumption of all goods over the life-cycle, so-called consumption smoothing, which indicate that periods with relatively low income result in increased borrowing. However, I will stand by the assumption that credit availability increase when households' net worth increase.

[15] In all three cases we assume that the purchase is (at least partly) debt financed.

[16] One may argue that existing dwellings become an inferior good when households' net worth increases sufficiently, as households then seek to "move up" in the housing market by entering the market for new dwellings, which can be seen as a form of quality improvement. Alternatively, one could also argue that housing is a normal good, as households buy a second home when net worth increases.

entering household due to it having serviced its loan for a longer period. In the latter case, if one household wants to "move up" in the housing market, i.e. buying a higher-quality and more expensive dwelling, another household downgrade correspondingly, i.e. buying a lower-quality and cheaper dwelling. The household which is buying the more expensive dwelling finances the difference with credit, leading to an increase in debt; the other party will have funds freed up, which, if used for other purposes than debt repayment, will increase total debt (however, the debt level will be unchanged if the funds are used to debt-servicing in its entirety). Hence, an increase in the housing turnover will leave households' total debt either unchanged or increased.

In the same manner as the housing price equation, equation (2.3.2), was transformed into a semi-logarithmic form, the "linearized" debt equation becomes,

$$d_t = \beta_{2,1} h_t + \beta_{2,2} yh_t + \beta_{2,3} R_t + \beta_{2,4} ph_t + \beta_{2,5} th_t \tag{2.4.2}$$

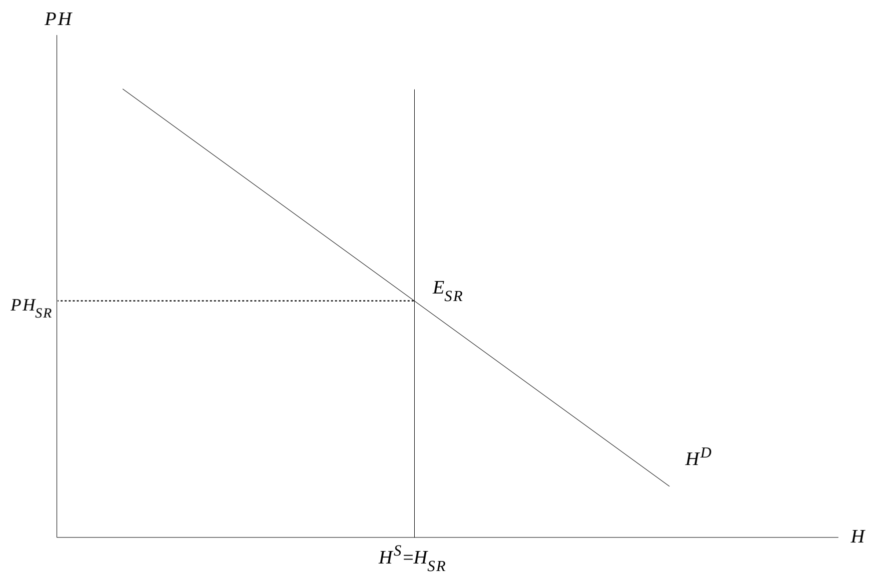


Figure 2: The housing market in the short-run

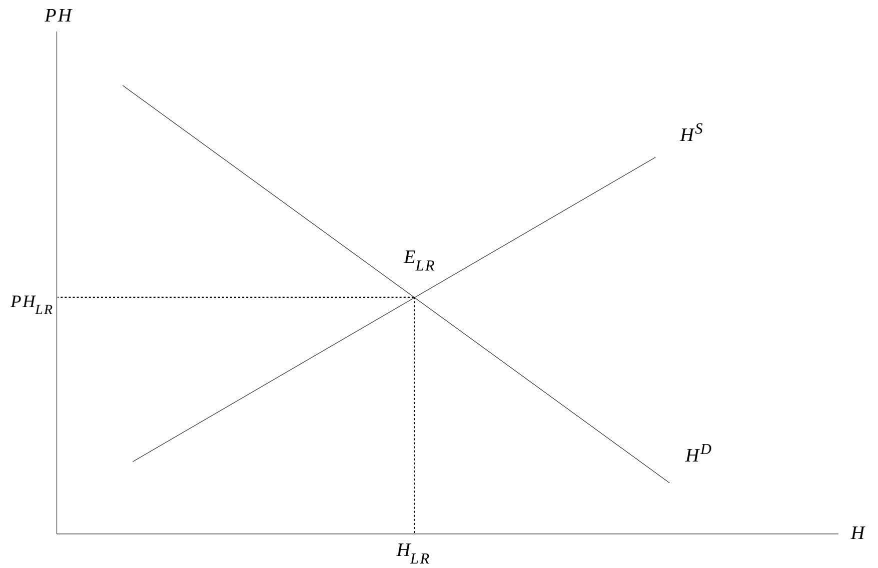


Figure 3: The housing market in the long-run

3 Data Description

The main objective of this thesis is to model econometrically price formation in three regional housing markets in Norway, namely Oslo & Akershus, the South-West which is composed of the counties Rogaland and Hordaland, and Northern-Norway which includes the three northernmost counties in Norway: Nordland, Troms and Finnmark. This chapter introduces and explains the main data series used in the analysis in chapters 6 and 7.

The regional data series used in chapter 7 come from county data, which have been aggregated to three regions with the use of unweighted arithmetic means. For instance, income data for the South-West is obtained by first collecting income data from Rogaland and Hordaland separately, and then combining them into one series by taking the mean. The rest of the data series used in the analysis are aggregate data on income, debt, housing stock, housing turnover, and the interest rate. All variables are quarterly and measured in real terms, i.e. they have been divided by the consumption deflator in the National Accounts. Finally, the variables have been transformed to *log*-scale, except for the real interest rate which is deflated by the CPI, and is kept on a linear scale. For the log-log elements in the model the accompanying coefficients are interpreted as elasticities, i.e. a one percent increase in the explanatory variable leads to a $\beta_{i,j}$ percentage increase in the dependent variable. Since, $R \approx log(1 + R)$ for small values of R, the coefficient on R_t can be treated as a semi-elasticity.

The aggregate variables are taken from the KVARTS database which is continously updated and revised by Statistics Norway. This is the same database Anundsen and Jansen (2013*b*) used in their analysis.

As the operational measure of housing prices, I use the housing price index provided by the Norwegian Association of Real Estate Agents (NEF),[17] which contains quarterly housing price data from 1990Q1, and monthly data from January 2002.[18] Furthermore, the housing price index is based on the price per square meter of the average sized dwelling of 100 m^2, and is a weighted average of three types of dwellings – single-unit dwellings, shared dwellings and apartments.

The interest rate series I have used is based on the average nominal lending rates of banks, where the real interest rate, $R = i(1 - \tau) - \pi$, is in line with the definition used in

[17]The dataset were handed to me by Samfunnsøkonomisk Analyse, an analyst agency.

[18]NEF's housing price index is a hedonic index based on detailed data of the houses' characteristics (Eitrheim and Erlandsen, 2004), and can be seen as the average price per unit of housing wealth. This is in contrast to the Case-Shiller index which uses the repeated-sales method.

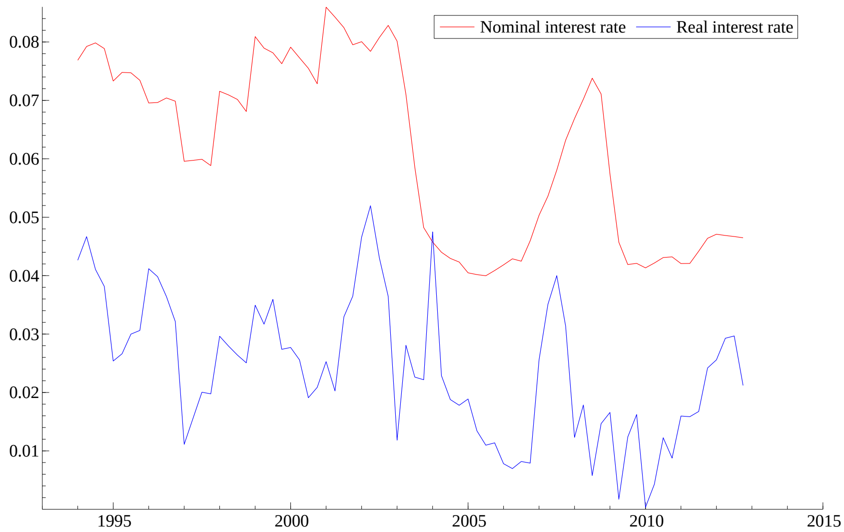


Figure 4: The nominal and real interest rate

Statistics Norway's macroeconomic models MODAG and KVARTS, cf. Boug and Dyvi (2008). Here, τ is the average marginal tax rate for employees, which has been held fixed at 28 percent since the tax reform in 1992, i is the nominal rate, and π is the consumer price index (CPI). The real after-tax interest rate is used as the operational measure of the user cost of housing, thus abstracting it from the expected real housing price appreciation which, instead, is captured by including lagged real price appreciation in the model. This is in line with similar analysis performed by Anundsen and Jansen (2013*b*), and it is consistent with Abraham and Hendershott's (1996) terminology of lagged housing prices acting as a "bubble builder" in the economy.

The rest of the data series are from Statistics Norway. The regional data on household gross income, wealth and debt are taken from the tax statistic for 2012, which consists of annual data available from 1993 to 2012.[19] Since these series are collected and reported annually, I have interpolated the series to get quarterly data. For the two stock variables; wealth and debt, I proceeded by simple linear interpolation, while I used an interpolation feature in *EViews* for the flow variable; income. Since it is unlikely that, e.g. quarterly income in any given year determines the household's choice of purchasing residential property (rather, it is reasonable to believe that it is the permanent income that determines such choices), such data manipulations are justified as it enable us to go forth with the analysis. The logs of these regional series on real income (yh), real wealth (w), and real debt (d), in addition to the real housing price series (ph), are presented in Figure 5.

The data series on regional housing prices are a bit fluctuating, but show a clear upward

[19]See Table 05661 for the former, and Table 05662 for the last two, in Statistics Norway's database: <ssb.no/statistikkbanken>

trend. The slump in prices in the start of the sample represents the Norwegian banking crisis in which housing prices plummeted. Since the nadir in 1992, housing prices reached their zenith in 2007, and eventually dropped markedly in the face of the financial turmoil of 2008. The downturn was, nevertheless, short lived as housing prices started to rise again in 2009, and has continued to grow ever since. Housing prices exert some regional differences, but the overall picture is that the regional housing markets, or rather that the regional housing price formation, are pretty synchronized and identical in terms of them having the same development trends. The series on income and wealth give the same overall impression. The data series on wealth might be affected by the change in the valuation of housing for tax purposes implemented in 2011. This is displayed by the kink in the graphs from 2010 to 2011, which is especially apparent in Oslo & Akershus and the South-West. Another interesting aspect is found by observing that in the time period from 1997 to 2003, when housing prices in Oslo & Akershus had a stronger growth rate than in the two other regions, also wealth in Oslo & Akershus was higher, which might indicate that part of the increase in wealth was associated with the increase in housing prices. There is also a kink in the income graphs in 2006, which may be explained by the tax reform introduced that year, which was the first tax reform implemented since 1992. The debt series exhibit a clear upward trend. This is in line with the evolution in income, wealth, and housing prices, e.g. as households net worth increases they can take on more debt, thus leading to higher housing prices and higher household wealth. This is illustrated nicely by looking at the income and debt series in the South-West in the last half of the sample; as households' income have increased relatively to Oslo & Akershus, effectively tightening the income gap between the two regions, the debt gap has also tightened. An obvious source of error in the data series originating from the tax statistics is incorrect reporting by respondents to the tax authorities. However, there is no reason to believe that systematic errors are present, and the data set is in this sense reliable. Naturally, the regional debt series are less fluctuating than the two other regional series originating from the tax statistics, as debt is less affected by changes in the tax rules due to the fact that debt is not subject to the incentives of misreporting regarding tax avoidance (to the same extent as income and wealth).

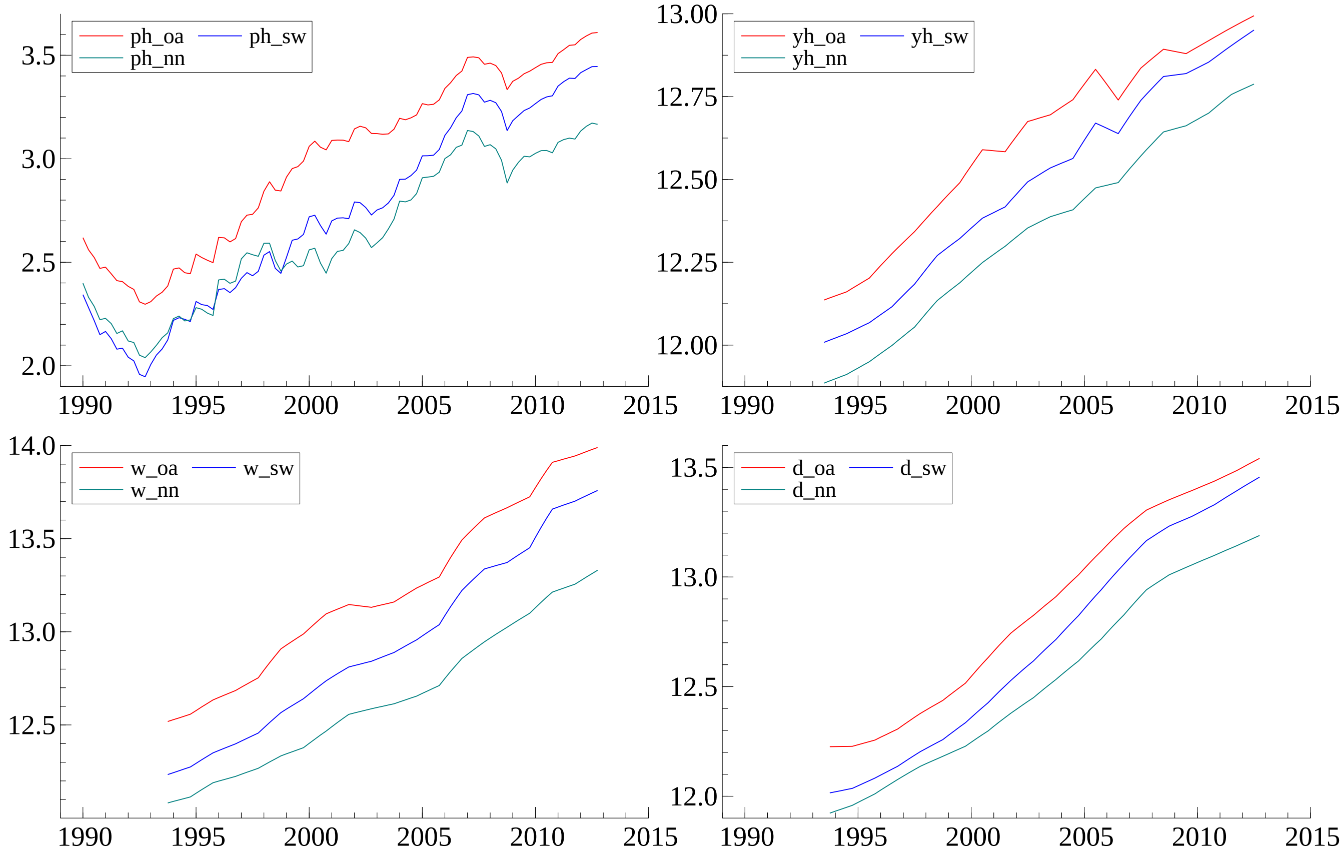


Figure 5: Graphical data description of the the regional variables used in the regional analysis.

4 Time-Series Econometrics

In this chapter I review some of the central modelling concepts and results from time-series econometrics that are used in chapters 6 and 7. While section 4.1 and 4.2 introduce concepts like stationarity and cointegration, and focus on why the models and the accompanying estimation results in chapters 6 and 7 can be given meaningful interpretation, section 4.3 is primarily concerned with giving the theoretical foundation of equilibrium correction models (ECMs) and standard cointegration tests, which is extensively used in the forthcoming analysis.

4.1 Stationarity

Stationarity is related to the linear properties of a time-series: expectation, variance and covariance. Formally, a time-series Y_t is stationary if it satisfies the following three stationarity conditions:

$$E(Y_t) = \mu \tag{4.1.1a}$$

$$var(Y_t) = \sigma^2 \tag{4.1.1b}$$

$$cov(Y_t, Y_{t+s}) = cov(Y_t, Y_{t-s}) = \gamma_s \tag{4.1.1c}$$

Thus, if the unconditional expectation, $E(Y_t)$, and the unconditional variance, $var(Y_t)$, exist, and are independent of t, and if the covariance, $cov(Y_t, Y_{t\pm s})$, is also, for any given s, independent of t, then our series is said to be *weakly stationary*, or *covariance stationary*. Consequently, if any of these conditions are violated, either by the first two moments or the covariance being time dependent, the series is nonstationary. If a series Y_t is stationary the empirical autocovariances will be consistent estimators for the theoretical autocovariances, and thus lay the foundation for consistent estimation of other parameters, e.g. coefficients in dynamic regression models (Bårdsen and Nymoen, 2014). In other words, stationarity is the main premise for why we can extend e.g. OLS based estimation and inference theory to time-series data.

Furthermore, there is a danger of obtaining apparently significant regression results from unrelated data when nonstationary series are used (Hill et al., 2008). Such regressions are said to be *spurious*, as emphasized by Granger and Newbold (1974). Hence, including nonstationary series in a regression model may result in it indicating a significant relationship when there is none. Since many macroeconomic time-series are nonstationary, one needs to be particularly cautious when estimating regressions with macroeconomic variables.

4.1.1 Order of Integration

Stationary series are said to be integrated of order zero, $I(0)$. Nonstationary series which can be made stationary by taking first differences are said to be integrated of order one, $I(1)$. Moreover, such series are said to contain a *unit root*, a term which we will return to below. In general, a series is integrated to order d, $I(d)$, if it must be differenced d times before the series becomes stationary; $(1-L)^d X_t \sim I(0)$, where L is the lag-operator.

Combining variables of the same integrating order in a regression produces a *balanced* regression, meaning that the variables we are studying have the same econometric properties. For instance, if both Y_t and X_t are $I(n)$, n being any integer greater than or equal to zero, then the regression

$$Y_t = \beta_1 + \beta_2 X_t + \varepsilon_t \tag{4.1.2}$$

is balanced. If they differed in their integrating order, e.g. Y_t being $I(0)$ and X_t being $I(1)$, then the regression would be *unbalanced*, and the results from OLS estimation cannot be given any meaningful interpretation (Granger, 1990). In the balanced case where both Y_t and X_t are $I(1)$, the corresponding residuals will in general also be $I(1)$, since in most cases a linear combination of a set of $I(1)$ variables will be $I(1)$ as well.

4.2 Cointegration

In general, one should be cautious when applying nonstationary series in a regression model due to the problem of obtaining spurious relationships. However, under the assumption that Y_t and X_t in equation (4.1.2) are $I(1)$ processes, there is a special case where it might exist one or more linear combinations of $I(1)$ variables that are $I(0)$. Such instances are examples of cointegration. The existence of a cointegrating relationship between $I(1)$ variables indicates that the series share similar stochastic trends, meaning that they satisfy one or more long-run relationships. Although they might diverge substantially from these relationships in the short-run, any deviations from these long-run relationships are temporary, as we will have an adjustment back to the "steady state". Hence, cointegration is the opposite of spurious regression. Because the error term can be expressed as $\varepsilon_t = Y_t - \beta_1 - \beta_2 X_t$, and is a stationary $I(0)$ process, the OLS estimators will be BLUE[20] in accordance with the Gauss-Markov theorem and standard inference theory is valid.

Formally, in line with Engle and Granger (1987), and refering to equation (4.1.2) in vector notation, the components of the vector $\mathbf{Y}_t$ are said to be cointegrated of order (d, b) denoted $\mathbf{Y}_t \sim CI(d, b)$, if

i) all components of $\mathbf{Y}_t$ are $I(d)$

ii) there exist a cointegrating vector $\boldsymbol{\beta}$ so that $\boldsymbol{\beta}'\mathbf{X}_t \sim I(d-b)$

Thus, two dependent $I(1)$ series which form a cointegrated relationship have the notation $CI(1, 1)$.

4.2.1 Unit Root Tests for Stationarity

Stationary series are characterized by having characteristic roots of the associated polynomial within the unit circle, i.e. eigenvalues less than unity in absolute value. Hence, one way to test for stationarity is to examine these characteristic roots. The most-widely used test for stationarity is the Dickey-Fuller test introduced by Dickey and Fuller (1979). Consider the following regression model as a starting point, where $Y_t \sim I(1)$ and $\varepsilon \sim N(0, \sigma^2)$

$$Y_t = \rho Y_{t-1} + \varepsilon_t \tag{4.2.1}$$

Next, the model represented by equation (4.2.1) can be reparametrized by subtracting Y_{t-1} from both sides

$$\Delta Y_t = (\rho - 1)Y_{t-1} + \varepsilon_t \tag{4.2.2}$$

If $\rho = 1$, then Y_t has a unit root and is a *random walk*. A random walk refers to the fact that the series is not drawn to any equilibrium value, but rather wanders slowly upwards or downwards, with no real pattern. When $\rho > 1$, the effect of the lags grow stronger and stronger with time, and Y_t display explosive behavior. In the case where $\rho < |1|$, the effects of earlier shocks "die out" with time, and the series is stationary. The null hypothesis in Dickey-Fuller test, $H_0 : \rho = 1$, corresponds to the series being nonstationary and having a unit root, while the alternative hypothesis, $H_1 : \rho < 1$, indicates stationarity. To take into account higher order dynamics (i.e. autocorrelation) an augmented Dickey-Fuller (ADF) test is conducted by including lags of the dependent variable, however, including

[20]Best Linear Unbiased Estimator.

too many lags will lead the test to lose power.[21] When the null hypothesis is true, the series is nonstationary and has a variance that increases as the sample size increases, thus, the distribution of the usual t-statistic is altered (see Hill et al., 2008, chap. 12.3.4). This means that, while the t-statistic still can be used, we must compare them to a set of critical values taken from the Dickey-Fuller distribution.[22] This distribution depends on whether equation (4.2.2) includes a constant, a trend or both (Davidson and MacKinnon, 2009). To recognize and distinguish it from the ordinary t-statistic the statistic in the Dickey-Fuller test is often called a τ-statistic.

The most important variable in the regional analysis presented in chapter 7 is the housing price variable. Thus, in order to verify the use of it as a stationary $I(1)$ series, a Dickey-Fuller unit root test for stationarity is conducted. As reported in Table 1, the test suggests that regional (real) housing prices follow an $I(1)$ process. The null hypothesis of nonstationarity is not rejected at a five percent significance level for any of the regions when examining the "raw" price series, e.g. as exemplified by Oslo & Akershus: $\tau^{OA} = -1.69 > -3.47 = \tau_c$. However, after taking first differences the test statistics come out significant, and the alternative hypothesis of stationarity is accepted; $\tau^{OA} = -6.15 < -2.90 = \tau_c$. Figure 6 shows the housing price series before and after taking first differences, alongside the autocorrelation function (ACF) which demonstrate the correlation between the residuals that are one period apart, two periods apart, and so on. Unit root tests for the rest of the variables used in the analysis are found in Appendix A.

Table 1: Unit root ADF test for the regional housing price series

Variabel	Lags	Constant	Trend	Seasonal	τ-ADF	5%-critical value
ph^{OA}	2	✓	✓	✓	-1.69	-3.47
ph^{SW}	2	✓	✓	✓	-2.47	-3.47
ph^{NN}	2	✓	✓	✓	-2.29	-3.47
Δph^{OA}	2	✓	–	–	-6.15	-2.90
Δph^{SW}	2	✓	–	–	-6.19	-2.90
Δph^{NN}	2	✓	–	–	-5.90	-2.90
Estimation period:		1994Q1-2012Q4				

[21] The power of a test is the probability that the test will reject the null hypothesis when the alternative hypothesis is true, and refers to the probability of *not* committing a type II error, i.e. failing to reject the null when it is false. So, when the power of a test is low it indicates a higher probability of making a type II error. A related concept is the level of significance, α, which is the probability of making a type I error, i.e. the probability of rejecting the null hypothesis when it is true.

[22] A detailed representation is given in Hamilton (1994).

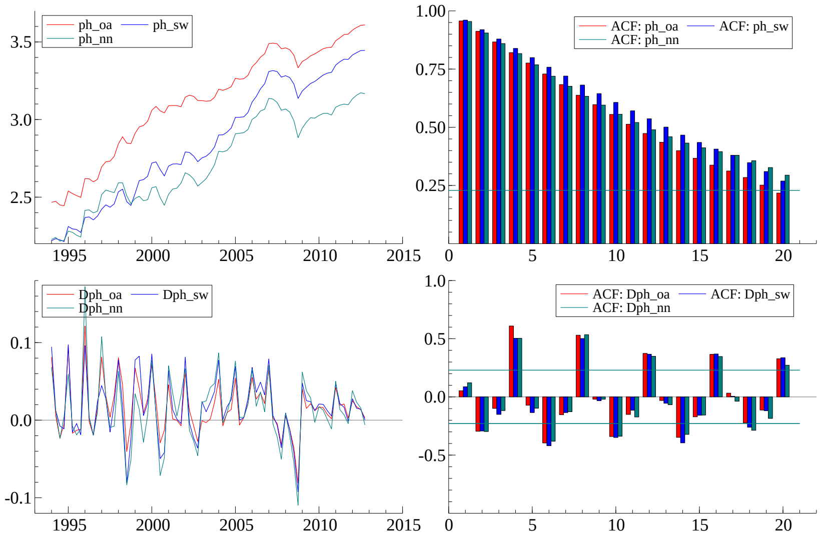


Figure 6: The stationarity properties of the real housing price series in the three regions.

4.3 Equilibrium Correction Models (ECM)

The Engle-Granger representation theorem states that cointegration implies equilibrium correction and vice versa (Engle and Granger, 1987). As mentioned above, cointegrated series tend to move together, and therefore they cannot drift "too" far apart, indicating that in the long-run the equilibrium steady state will be (re-)established. This cointegrated relationship can be represented as an equilibrium correction model (ECM), which will capture the short-run dynamics as well as the effects of deviations from the long-run equilibrium.

Consider the following autoregressive distributed lag (ARDL) model[23]:

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \beta_1 X_t + \beta_2 X_{t-1} + \varepsilon_t \tag{4.3.1}$$

where $\varepsilon_t \sim IID(0, \sigma^2)$. Equation (4.3.1) is an ARDL(1,1) model as it includes one lag of the dependent variable, Y_t, and one lag of the explanatory variable, X_t, and is a special case of the more general ARDL(p,q) model. Dynamic models like the ARDL provide us with a tool for analyzing the short-run and long-run dynamics of the system. The ECM representation is just a reparametrization of the ARDL model in equation (4.3.1), and it expresses equation (4.3.1) in terms of an error-correction mechanism:

$$\Delta Y_t = \phi_0 + \beta_0 \Delta X_t + (\phi_1 - 1)(Y_{t-1} - \frac{\phi_0}{1-\phi_1} - \frac{\beta_0 + \beta_1}{1-\phi_1} X_{t-1}) + \varepsilon_t \tag{4.3.2}$$

[23]In the literature the ARDL is often abbreviated ADL.

$$\Delta Y_t = \phi_0 + \beta_0 \Delta X_t + (\phi_1 - 1)(Y_{t-1} - \mu^*_{Y|X_{t-1}}) + \varepsilon_t \tag{4.3.3}$$

where $\mu^*_{Y|X_{t-1}}$ is the conditional (long-run) equilibrium of Y_t given X_{t-1}, defined as

$$\mu^*_{Y|X_{t-1}} = \frac{\phi_0}{1-\phi_1} + \frac{\beta_0+\beta_1}{1-\phi_1} X_{t-1} \tag{4.3.4}$$

$\frac{\beta_0+\beta_1}{1-\phi_1}$ is the long-run multiplier, and it is an elasticity if the model is on log-log form. It is clear that the short-run and long-run features of the dynamic relationship is modelled separately in the ECM framework, where β_0 describes the short-run relationship of a change in X_t on Y_t, while $(\phi_1 - 1)$ captures the speed of adjustment to equilibrium. In other words, if the system is outside its "steady state," e.g. after being exposed to a shock, $y_{t-1} \lesseqgtr \mu^*_{Y|X_{t-1}}$, the equilibrium correction term indicate how fast the system will adjust back towards the equilibrium by increasing or decreasing ΔY_t depending on whether the system is below or above $\mu^*_{Y|X_{t-1}}$ respectively. Naturally, any deviation from equilibrium $y_{t-1} \neq \mu^*_{Y|X_{t-1}}$, must result in a negative error correction coefficient, $(\phi_1 - 1) < 0$, in order for the system to return to equilibrium. Thus, in the ECM represented by equation (4.3.3), changes in the endogenous variable, ΔY_t, can be explained by two factors:

- Changes in the exogenous variable, ΔX_t, or
- Equilibrium correction of last period's deviation from the long-run equilibrium given by $(Y_{t-1} - \mu^*_{YX_{t-1}})$

4.3.1 Johansen Trace Test

Consider the following special case of an ARDL, namely the Gaussian VAR of the first order

$$\mathbf{Y}_t = \mathbf{\Phi} \mathbf{Y}_{t-1} + \boldsymbol{\varepsilon}_t \tag{4.3.5}$$

The vector $\mathbf{Y}_t$ consists of $n \times 1$ variables,[24] $\mathbf{\Phi}$ is the $n \times n$-coefficient matrix, and $\boldsymbol{\varepsilon}_t$ is a vector consisting of normally distributed disturbances. Equation (4.3.5) can be rewritten as

$$\Delta \mathbf{Y}_t = \mathbf{\Phi}^* \Delta \mathbf{Y}_{t-1} + \mathbf{\Pi} \mathbf{Y}_{t-1} + \boldsymbol{\varepsilon}_t \tag{4.3.6}$$

[24] If $n = 3$ and we only include on lag of the dependent variable, then we have a $(n = 3)$-dimensional VAR of the first order, with $\mathbf{Y}_t = (Y_t, X_t, Z_t)'$.

Here, $\mathbf{\Pi} = \mathbf{\Phi} - \mathbf{I} = \boldsymbol{\alpha}\boldsymbol{\beta}'$ represents the $n \times n$-levels-coefficient-matrix of the lagged variables, where $\boldsymbol{\alpha}_{n\times r}$ is the vector of the adjustment coefficients, $\boldsymbol{\beta}_{r\times n}$ is the vector of cointegration coefficients, and r refers to the rank of $\mathbf{\Pi}$.

The Johansen trace test is all about testing the rank of the $\mathbf{\Pi}$-matrix in equation (4.3.6). Because the rank of a matrix is related to the number of independent linear combinations of the matrix, the rank of $\mathbf{\Pi}$ will correspond to the number of cointegrating relationships in the VAR. Thus, since the rank of $\mathbf{\Pi}$ is given by the number of non-zero eigenvalues of $\mathbf{\Pi}$, one approach to testing for cointegration is to find the number of eigenvalues that are significantly different from zero and less than unity in absolute value.

The trace test is a *sequential* testing procedure which starts by testing the hypothesis of $r \leq 0$, i.e. rank (less than, or) equal to zero, against the alternative $r \geq 1$, i.e. rank greater than or equal to one. If the null is rejected, we proceed by testing $r \leq 1$ against $r \geq 2$, and continue in this fashion until the conclusion is non-rejection of the null hypothesis. Thus, the procedure yields the conclusion that there are $r + 1$ cointegrating vectors if the last significant test is η_r. Notice that if we reject the test with the null hypothesis of $r \leq (n-1)$, then $\mathbf{\Pi}$ has full rank, $r = n$, and all eigenvalues are significantly different from zero, and hence the VAR is weakly stationary; i.e. VAR $\sim I(0)$.

A further remark concerns the critical values obtained in the trace test. As with the Dickey-Fuller distributions in the unit root tests, the distributions used to obtain critical values in the Johansen procedure are also non-standard. For instance, the distributions depend on which deterministic variables are included in the model, and whether or not they are included restricted or unrestricted. Doornik (2003) provides some details.

After having determined the cointegrating rank one can formulate a *cointegrated VAR*[25] which is then a stationary dynamic system. Further, one can obtain an identified simultaneous equation system (SEM) by the order and rank condition by imposing identifying restrictions based on relevant economic theory. Moreover, one can also impose and test overidentifying restrictions on the system.

The Johansen procedure is a full-fledged method for testing the cointegrating relationships of a system. This is in contrast to the Engle-Granger and ECM test which can only be conducted if there are no more than one cointegrating relationship. Thus, when there are more than one cointegrating relationship, meaning that there are more than one variable that equilibrium corrects, the Johansen trace test is the appropriate procedure to use, as the underlying assumption of weak exogeneity in the two alternative tests does not hold. Weak exogeneity means that statistically efficient estimation and inference can be

[25]In the literature one also comes across the term vector equilibrium correction, or VECM.

achieved by only considering the conditional model without taking the rest of the system into account, i.e. abstracting it from the marginal model(s).

4.3.2 ECM Test

As mentioned briefly in the paragraph above, there are two other methods to test for cointegration which can be applied when there is only one cointegrating relationship, namely the Engle-Granger and ECM test. The latter is related to the Engle-Granger representation theorem stating that cointegration implies equilibrium correction. With equation (4.3.3) as our ECM, we can test the hypothesis of no cointegration, $H_0 : \psi = (\phi_1 - 1) = 0$, against the alternative of cointegration, $H_1 : \psi = (\phi_1 - 1) < 0$. This is because a non-zero ψ-coefficient indicates an error correction mechanism. The corresponding ECM test statistic will not follow the standard normal, $N(0, 1)$, distribution asymptotically, rather it will follow a $\kappa_d(g)$-distribution, where d indicate if we are looking at a case with *nc*, *c*, *ct*, or *ctt*,[26] and g is, if Y_t in equation (4.3.3) is to be treated as a vector, the number of columns. In the special case where $g = 1$, the asymptotic distribution of the ECM statistic is identical to that of the corresponding Dickey-Fuller τ statistic.

Because the error-correction term often has considerably explanatory power when there is cointegration it is less likely to suffer from serial correlation than the Engle-Granger test (Davidson and MacKinnon, 2009). When there is only one cointegration vector and $n - 1$ weakly-exogenous variables the ECM test is just a special case of the Johansen procedure. In the econometric jargon; when the cointegrating relationship appear only in the equation of interest, the Johansen procedure reduces to estimation based on a single ECM equation and OLS, thus neglecting the marginal model of the system in line with the weak exogeneity assumption.Thus, the ECM test (and the tests based on there only being one cointegrating relationship) depends crucially on the assumption of weak exogeneity of the exogenous variables of the system.

4.3.3 Engle-Granger Test

The Engle-Granger (EG) test is the simplest way to test for cointegration, and it is conducted in two-steps. The first step consists of obtaining the residuals from the cointegration regression, e.g. equation (4.1.2). Under the assumption that both Y_t and X_t are cointegrated $I(1)$ variables, the error term will be $\varepsilon_t \sim I(0)$. The second step is then to

[26] Abbreviations for no constant, constant, constant and trend, and constant and trend squared. This is relevant in the same way as it was in the Dickey-Fuller unit root tests for stationarity, namely through the fact that it determines the distribution of the test statistic.

subject these estimates, $\hat{\varepsilon}_t$, to an ADF test, by running the regression

$$\Delta\hat{\varepsilon}_t = \theta\mathbf{X}_t + \gamma\hat{\varepsilon}_{t-1} + \sum_{j=1}^{p}\delta_j\Delta\hat{\varepsilon}_{t-j} + e_t \tag{4.3.7}$$

where $\mathbf{X}_t$ contains a constant term, and p is the number of lags we choose to include in the regression to correct for residual autocorrelation. Thus, the ADF test is performed under the null hypothesis of no cointegration, $H_0 : \gamma = 0$, and are assessed against a set of critical τ-values which are higher than in the regular ADF test because we already have estimated the regression coefficient. Put differently, if the residuals are nonstationary, then the series are cointegrated, while stationary residuals are a sign of cointegrated series. A precautionary note regarding the power of the EG test should nevertheless be taken: Compared to the ECM test the power of the EG test is in general lower, meaning that the probability of a type II error is greater, i.e. failing to reject the null hypothesis of no cointegration when it is in fact wrong. Furthermore, including more explanatory variables in the model, (4.1.2), will make the critical values even higher in absolute value. In consequence, the power of the test diminishes. Additionally, we observe that the EG test can be interpreted as a *restricted* ECM test.[27]

[27]See Bårdsen and Nymoen (2014, chap. 11.4.3) for a disposition in Norwegian.

5 Literature Review

As the housing market has become an integral part in the economy in most developed countries, it has been a central topic of research. The purpose of this chapter is to introduce some of the lines of research within the branch of housing economics.

The literature on housing price formation is extensive, see e.g. Hendry (1984) and Muellbauer and Murphy (1997) for a theoretical disposition, and Girouard et al. (2006) for an overview of the empirical literature. Anundsen (2013) points out that there are generally two different theoretical approaches that are considered when looking at the relationship between housing prices and fundamentals; the inverted demand approach and the price-to-rent approach.[28]

As emphasized by Gallin (2006), Mikhed and Zemčík (2009), and Anundsen (2013), there is no consensus in the literature with regards to whether housing prices and fundamentals are cointegrated or not, and there is also a dichotomy between those who use aggregate time-series data, and those who use regional data.[29] There is relatively little formal evidence of cointegration, but some researchers are, nevertheless, able to conclude that housing prices and fundamentals are cointegrated based on test statistics that are found to be close to their critical values.

When modelling housing prices the assumption that there exists a cointegrating relationship between housing prices and fundamentals is important as it has major implications for how we set up the econometric model, as well as for the corresponding interpretations and conclusions. Thus, if prices and income are cointegrated, an error correction specification is appropriate as the gap between the two will be a useful indicator of whether housing prices are above or below the equilibrium "steady state" values. On the other hand, if there does not exist any cointegrating relationship between the two, then such error correction models widely applied in the literature are inappropriate, and adjustment towards an equilibrium need not happen, effectively implying that housing prices can deviate substantially from fundamentals (Gallin, 2006).

In addition, as Flood and Hodrick (1990) draw attention to, it is both important and difficult to distinguish between fundamentals-driven housing price changes and bubbles, and it is wise to avoid making bombastic statements. After all, price movements that are interpreted as bubbles could even be a product of model mis-specification since we can never know the "true" model.

[28]Both these approaches are used in a life-cycle framework for modelling housing prices. See Anundsen (2013) for a thorough discussion.

[29]Regional data are sometimes referred to as micro data or panel data in the literature.

An important panel data study by Abraham and Hendershott (1996), investigate housing prices based on annual data on 30 Metropolitan Statistical Areas (MSA) in the U.S. for the period 1977–1992. In their equilibrium correction type of model they find evidence for that income, the real interest rate, and construction costs are important determinants for explaining housing prices in the long-run. In their terminology, lagged housing price appreciation are included in the model representing a "bubble builder" effect by capturing the momentum effect: as, say, an increase in fundamentals leads to higher housing prices, the lagged component in the model will magnify this housing price increase. Furthermore, the "bubble burster" is proxied by the adjustment coefficient in the ECM framework; if the adjustment coefficient is close to (or equal to) zero, deviations from the an estimated equilibrium would be restored very slowly – or not at all.

Two central works that have attracted some interest are Malpezzi (1999) and Duca et al. (2011). Malpezzi (1999) employs an error correction specification to estimate how "effective" the housing market is, i.e. an effective well-functioning market being characterized as one where prices remain stable as increases in demand implies increases in the supply of housing. Using data for 133 metropolitan areas over the years 1979–1996, Malpezzi finds evidence that the determinants for explaining housing prices are per capita income and its growth rate, the nominal mortgage interest rate, population and its growth rate, in addition to the regulatory environment. On the other hand, Duca et al. (2011) applies a price-to-rent framework, and find that credit standards, as measured by the average loan-to-value (LTV) ratio for first-time home-buyers affect housing prices. This finding supports the view that the easing of the U.S. mortgage standards in the early 2000's were effectively raising demand for housing, and that the corresponding reversal of some of these practices led to the turmoil associated with the U.S. subprime mortgage crisis.

The development in the housing markets can also be explained in a framework incorporating a Tobin-q effect (Tobin, 1969). As reported in Sørensen and Whitta-Jacobsen (2010), a q-theory of housing investment is based on the fact that rising market prices for existing residential property relative to the cost of construction of a new unit of housing makes it more profitable for firms in the construction industry to increase the supply of housing.

Studies on the Norwegian housing markets have been conducted by Eitrheim (1993), Jacobsen and Naug (2004*a*), and Anundsen and Jansen (2013*b*). Eitrheim (1993) focuses on the housing price equation that formed the basis for RIMINI, Norges Bank's former macroeconomic model, and concludes that disposable income, real gross debt, and the value of the housing stock are the most important determinants for explaining housing prices. Moreover, the short-run dynamics include additional factors such as the inflation rate, the unemployment rate, capital tax and lending rates.

In Boug and Dyvi (2008), an explanation of the housing sector in MODAG[30] is presented. As commented on in chapter 2, this framework forms the basis for Anundsen and Jansen (2013*b*), as well as my own model which is presented in chapter 7.

Jacobsen and Naug (2004*a*) investigate further which factors drives the housing prices in Norway, using a sample of quarterly data over the period 1990 to 2004. They find that the interest rate, new construction, unemployment and households' income are the most significant factors that explain housing price developments. They also find that housing prices equilibrium correct by 12 percent each quarter when housing prices deviate from their steady state level. In a related paper, Jacobsen and Naug (2004*b*), it is concluded that the debt increases among Norwegian households over the same interval is connected with the developments in the housing market. This anticipates Anundsen and Jansen (2013*b*) study of self-reinforcing effects between housing prices and credit, where the authors jointly estimate the long-run interactions between housing prices and household debt for a post-deregulation sample, 1984–2008, using quarterly data. They find that housing prices depend on household borrowing, real disposable income and the housing stock in the long-run, whereas real household debt is driven by the value of housing stock,[31] the real interest rate, and the housing turnover.[32] The equilibrium correction coefficient associated with Anundsen and Jansen (2013*b*) is 17.5 percent, indicating that housing prices are expected to correct by 0.175 percentage points for every percentage point of disequiliubrium in housing prices each quarter. In other words it will take a little less than six quarters for housing prices to return to equilibrium.

[30]Statistics Norway's macroeconomic models MODAG and KVARTS are more or less identical, with the main difference being that while former uses annual data, the latter employs quarterly data.

[31]Which is defined as housing prices multiplied with the housing stock.

[32]These parsimonious models emerge after imposing testable overidentifying restrictions on the system. In the re-estimation part in chapter 6 these restrictions are introduced and explained.

6 Replication of the Housing Price Model in Anundsen and Jansen (2013)

Anundsen and Jansen (2013*b*), henceforth A&J (2013), study the self-reinforcing effects between housing prices and credit, and in this chapter their model is re-estimated with the macroeconomic KVARTS database which is developed and maintained by Statistics Norway. Because of data revisions there are some differences between this dataset and the original dataset used by A&J (2013).

The purpose of my replication and re-estimation is twofold. First, it is of interest in itself to check the robustness of the results in A&J (2013) with respect to the changes in the measurement, and the degree of parameter stability with respect to the extension of the sample period. Second, if the results still hold and are robust, it could provide a reference point for my modelling of regional housing prices in Norway.

The sample period used in A&J (2013) is 1986Q4–2008Q4, and it is in practice a post-deregulation sample. A&J argue that extending the sample backwards would make the model less suitable to give answer about the self-reinforcing relationship between housing prices and credit due to the regulation of the credit market in earlier periods, which clearly distorted the ordinary market mechanisms.

6.1 Re-Estimation of Anundsen and Jansen (2013)

The Johansen trace test is used to decide on the number of cointegrating relationships. By applying the 5%-critical values in A&J (2013) I find evidence for one cointegrating relationship, see Table 2. This is in contrast to the results in A&J (2013) which find evidence for two cointegrating relationship.[33] However, by looking closely at the results in A&J (2013) it is clear that the null hypothesis of (weakly less than) one cointegrating relationship is barely rejected in favor for the alternative hypothesis of more than one cointegrating relationship. Thus, these minor (yet, potentially important) deviations, could be attributed the fact that I used an updated and revised dataset relative to A&J (2013).

Moreover, it is reassuring to observe that the standard test battery yield satisfying results. The AR test indicates non-rejection of the null hypothesis of no autocorrelation based on the p-value.[34] The null hypothesis of the normality test is that of normally distributed

[33]The critical values are obtained from Table 13 in Doornik (2003) with three exogenous variables.

[34]Reject the null hypothesis when the p-value is less than, or equal to, the level of significance. That

disturbances. The test is insignificant which leads to non-rejection of the null hypothesis, and, therefore, inference based on the t- and F-distributions is correct. The hetero test is also insignificant, indicating that the null hypothesis of homoskedastic residuals is not rejected. Since all the mis-specification tests are insignificant, the formal evidence suggests that there are not enough evidence in the data to reject the null hypothesis of no mis-specification, hence, corresponding estimation and inference is valid. A more thorough description of the mis-specification tests is found in Appendix B.

Table 2: Trace test for cointegration over the sample 1986Q2–2008Q4

Eigenvalue: λ_i	H_0	H_A	λ_{trace}	5%-critical value
0.41	$r = 0$	$r \geq 1$	84.16	64.48
0.21	$r \leq 1$	$r \geq 2$	35.38	40.95
0.14	$r \leq 2$	$r \geq 3$	13.86	20.89
Diagnostics	Test statistic	Value (p-value)		
Vector AR 1-5 test:	$F(45{,}146)$	0.91 (0.63)		
Vector Normality test:	$\chi^2(6)$	7.93 (0.24)		
Vector Hetero test:	$F(270{,}247)$	0.95 (0.66)		
Estimation period:	1986Q2-2008Q4			

Accordingly, the authors impose a set of (untestable) identifying restrictions in order to get exact identification of the system. Normalizing on real housing prices and household debt in the two respective relations, then assuming that the housing turnover has no direct effect on real housing prices, and that a change in either the housing stock or housing prices have the same effect on household debt, leaves us with a just-identified system consisting of housing price- and debt equation.

$$ph = \beta_{d,1}d + \beta_{yh,1}yh + \beta_{h,1}h + \beta_{R,1}R + \beta_{t,1}t \tag{6.1.1}$$

$$d = \beta_{ph,2}ph + \beta_{yh,2}yh + \beta_{R,2}R + \beta_{th,2}th + \beta_{h,2}h + \beta_{t,2}t \tag{6.1.2}$$

Even though I did not find formal evidence for two cointegrating relationships, I will proceed under the assumption that the rank is two so that I can follow the analysis by A&J (2013) as closely as possible. The replication is continued in Table 3 where overidentifying restrictions are imposed and tested. We observe that the replication of Panel 1, in which we test for no trend in equation (6.1.1) and (6.1.2), shows coefficient estimates that deviate somewhat from those in the published paper. However, with a

is, if $p \leq \alpha$, then reject H_0. Thus, the p-value indicates the marginal significance level associated with the test statistic, i.e. the greatest level for which a test fails to reject the null.

standard error of 0.81 it is also clear that the interest rate in the housing price equation is insignificant, which leads me – just as in A&J's discussion paper – to Panel 2, where we impose the restriction of no effect of real after tax interest rate on housing prices. In the same way as in the first panel, we still have quite inflated values of the real housing stock in the housing price equation – it is more than twice as large in magnitude – and the real interest rate and housing turnover in the debt equation – they are, respectively, over ten times and eight times as big as those values reported in A&J's analysis. Nevertheless, it is encouraging to see that the signs of the coefficients are correct.

In Panel 3, in which we assume no effect of disequilibrium housing prices on household debt, our estimates of the debt equation is quite satisfying in terms of replication. As is natural, parameters become more precisely estimated when we impose more restrictions on the system. Panel 4 gives the same overall impression, namely that some estimates are a bit inflated in magnitude, but that the signs are correct.

In Panel 5 we impose weak exogeneity of income with respect to the long-run coefficients. The coefficients have the correct signs and compare well to the estimation results in the discussion paper. The reader is asked to confer with Appendix C for the corresponding estimation results in A&J (2013*b*). Furthermore, the equilibrium correction coefficients are almost identical to A&J (2013), and they are also very significant, e.g. a t-value greater than 2.90. Since all the diagonal elements in the $\boldsymbol{\alpha}$-matrix[35] are significant with low p-values, this would indicate that $r = 2$, i.e. two cointegrating relationships. We also note that the second eigenvalue, $\lambda_2 = 0.21$, is significantly different from zero. Furthermore, the χ^2-tests reported in Table 3 are the LR test of overidentifying restrictions,[36] which conclude that our model is correctly specified, indicating that the imposed restrictions on the system are valid, as we do not reject the null hypothesis of correct specification at any conventional significance level. The estimated error correction terms in Figure 7 indicate that the period associated with the Norwegian banking crisis, 1988–1993, indeed was characterized by a significant amount of noise in the housing and credit markets, while the last part of the sample suggests that the markets have been more or less in steady state as reflected by the stationarity of the series.

The conclusion is that that the results in A&J (2013) are more or less robust to my re-analysis, even with a dataset that has been changed and revised. These results are reassuring, and provide us with useful knowledge and a solid foundation in the continuation of the housing price analysis.

[35] Recall from chapter 3 that the $\boldsymbol{\alpha}$-matrix is the matrix consisting of the adjustment coefficients.

[36] This test is also known as the Sargan test, Hansen-Sargan test, and J-test.

Table 3: Testing steady-state hypotheses over the sample 1986Q2–2008Q4

The just identified housing price and debt equation are defined by

$$ph = \beta_{d,1}d + \beta_{yh,1}yh + \beta_{h,1}h + \beta_{R,1}R + \beta_{t,1}t$$

$$d = \beta_{ph,2}ph + \beta_{yh,2}yh + \beta_{R,2}R + \beta_{th,2}th + \beta_{h,2}h + \beta_{t,2}t$$

Panel 1: Testing no trend ($\beta_{t,1} = \beta_{t,2} = 0$)

$$ph = \underset{(0.23)}{1.03}d + \underset{(0.84)}{2.81}yh - \underset{(1.54)}{4.40}h - \underset{(0.81)}{0.03}R$$

$$d = 0.71ph - \underset{(1.48)}{1.43}yh - \underset{(3.95)}{7.88}R + \underset{(0.45)}{0.96}th + \underset{(0.51)}{0.71}h$$

$$LogL = 850.880,\ \chi^2(2) = 3.75(0.15)$$

Panel 2: No effect of real after tax interest rate on house prices ($\beta_{R,1} = 0$)

$$ph = \underset{(0.23)}{1.04}d + \underset{(0.83)}{2.82}yh - \underset{(1.54)}{4.42}h$$

$$d = 0.71ph - \underset{(1.48)}{1.43}yh - \underset{(3.95)}{7.88}R + \underset{(0.45)}{0.96}th + \underset{(0.51)}{0.71}h$$

$$LogL = 850.880,\ \chi^2(3) = 3.75(0.29)$$

Panel 3: No effect of disequilibrium housing prices in household debt

$$ph = \underset{(0.22)}{1.02}d + \underset{(0.81)}{2.76}yh - \underset{(1.48)}{4.30}h$$

$$d = 1.19ph - \underset{(0.81)}{1.73}yh - \underset{(2.15)}{4.09}R + \underset{(0.25)}{0.58}th + \underset{(0.28)}{1.19}h$$

$$LogL = 850.743,\ \chi^2(4) = 4.02(0.40)$$

Panel 4: No effect of real disposable income on household debt ($\beta_{yh,2} = 0$)

$$ph = \underset{(0.23)}{1.00}d + \underset{(0.81)}{2.44}yh - \underset{(1.49)}{3.84}h$$

$$d = 0.71ph - \underset{(1.82)}{3.51}R + \underset{(0.14)}{0.17}th + \underset{(0.14)}{0.71}h$$

$$LogL = 848.421,\ \chi^2(5) = 8.67(0.12)$$

Panel 5: Imposing weak exogeneity of income with respect to the long-run coefficients

$$ph = \underset{(0.23)}{1.04}d + \underset{(0.82)}{2.63}yh - \underset{(1.51)}{4.23}h$$

$$d = 0.74ph - \underset{(2.16)}{3.91}R + \underset{(0.16)}{0.21}th + \underset{(0.16)}{0.74}h$$

$$\alpha_{1,ph} = \underset{(0.03)}{-0.22},\ \alpha_{1,d} = \underset{(0.02)}{-0.07},\ \alpha_{2,d} = \underset{(0.01)}{-0.04}$$

$$LogL = 847.165,\ \chi^2(7) = 11.18(0.13)$$

The sample is 1986Q2 to 2008Q4, 91 observations

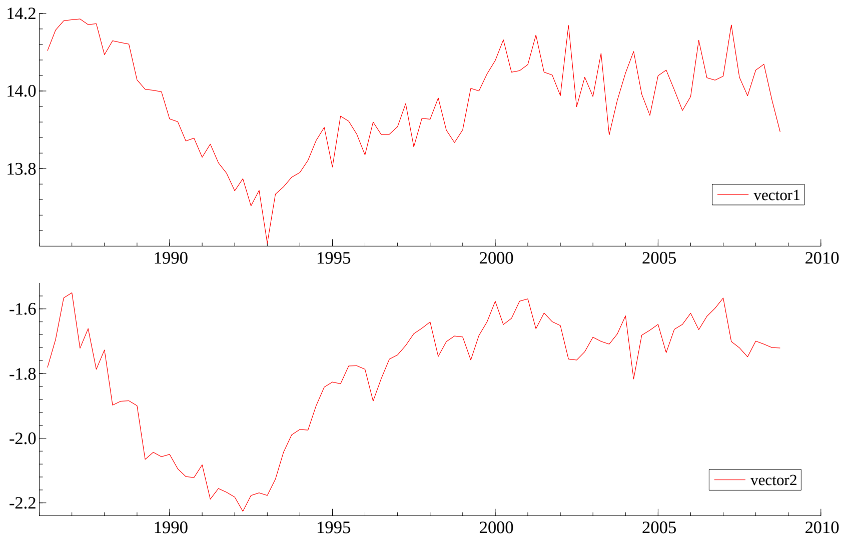


Figure 7: The estimated error correction terms of the two cointegrating vectors associated with the re-estimation of Panel 5

6.2 Estimation of Anundsen and Jansen (2013) over an Extended Sample

A&J's (2013) sample spans over a post-deregulation period and ends in the fourth quarter of 2008, which coincides with the outbreak of the financial crisis. By extending the data sample, and incorporating the first four years of the crisis, we get a better picture of how robust the results of A&J (2013) are.

Table 4 shows the trace test results for the extended sample, and it is clear that it indicates proof for two cointegrating vectors. However, a low p-value is observed for the normality test, the null hypothesis of normally distributed residuals is rejected at $\alpha_{SL} = 0.05$, indicating that the model does not fit the data as good over the extended sample. This suggests that the last observations in the sample is characterized by "noisy" observations, i.e. statistical outliers, associated with the latest financial crisis, which affects the model. The estimation results when imposing overidentifying restrictions on the cointegrated VAR are presented in Table 5. Even though there are some less encouraging signs on ph, yh and h in the debt-equation in Panel 1 and 2, they are not to be taken too seriously as they are clearly insignificant, and they have the expected sign after imposing more restrictions on the system. The coefficient estimates are more or less comparable to the results in the previous part of the analysis, suggesting that our results are robust, especially taking into consideration that the last 16 observations of the estimation period have coincided with the turmoil associated with the financial crisis' outbreak in 2008. In other words, my results suggest parameter stability as the coefficient estimates are not subject to any significant changes when we extend the sample period, which add depth to the conclusion in the previous section that the results in A&J (2013) provide a good starting point, and acts as a benchmark for further analysis of housing prices.

Table 4: Trace test for cointegration over an extended sample, 1986Q2–2012Q4

Eigenvalue: λ_i	H_0	H_A	λ_{trace}	5%-critical value
0.36	$r = 0$	$r \geq 1$	89.57	64.48
0.20	$r \leq 1$	$r \geq 2$	42.11	40.95
0.16	$r \leq 2$	$r \geq 3$	18.11	20.89
Diagnostics	Test statistic	Value (p-value)		
Vector AR 1-5 test:	F(45,193)	1.13 (0.29)		
Vector Normality test:	$\chi^2(6)$	16.46 (0.01)*		
Vector Hetero test:	F(270,342)	1.09 (0.22)		
Estimation period:	1986Q2-2012Q4			

Table 5: Testing steady-state hypotheses over an extended sample, 1986Q2–2012Q4

The just identified housing price and debt equation are defined by

$$ph = \beta_{d,1}d + \beta_{yh,1}yh + \beta_{h,1}h + \beta_{R,1}R + \beta_{t,1}t$$
$$d = \beta_{ph,2}ph + \beta_{yh,2}yh + \beta_{R,2}R + \beta_{th,2}th + \beta_{h,2}h + \beta_{t,2}t$$

Panel 1: Testing no trend ($\beta_{t,1} = \beta_{t,2} = 0$)

$$ph = \underset{(0.22)}{0.98}d + \underset{(0.75)}{3.51}yh - \underset{(1.40)}{5.03}h - \underset{(1.07)}{2.03}R$$
$$d = -3.01ph + \underset{(3.99)}{5.36}yh - \underset{(14.17)}{35.26}R + \underset{(1.10)}{2.17}th - \underset{(1.45)}{3.01}h$$
$$LogL = 990.504,\ \chi^2(2) = 2.84(0.24)$$

Panel 2: No effect of real after tax interest rate on house prices ($\beta_{R,1} = 0$)

$$ph = \underset{(0.21)}{1.43}d + \underset{(0.84)}{4.24}yh - \underset{(1.51)}{7.07}h$$
$$d = -1.27ph + \underset{(3.96)}{1.94}yh - \underset{(9.54)}{22.95}R + \underset{(1.09)}{1.89}th - \underset{(1.34)}{1.27}h$$
$$LogL = 989.727,\ \chi^2(3) = 4.39(0.22)$$

Panel 3: No effect of disequilibrium housing prices in household debt

$$ph = \underset{(0.22)}{1.34}d + \underset{(0.85)}{4.06}yh - \underset{(1.54)}{6.60}h$$
$$d = 1.06ph - \underset{(0.81)}{1.55}yh - \underset{(1.94)}{4.26}R + \underset{(0.22)}{0.62}th + \underset{(0.27)}{1.06}h$$
$$LogL = 988.933,\ \chi^2(4) = 5.98(0.20)$$

Panel 4: No effect of real disposable income on household debt ($\beta_{yh,2} = 0$)

$$ph = \underset{(0.21)}{1.35}d + \underset{(0.81)}{3.43}yh - \underset{(5.78)}{5.78}h$$
$$d = 0.63ph - \underset{(1.48)}{3.40}R + \underset{(0.11)}{0.22}th + \underset{(0.09)}{0.63}h$$
$$LogL = 987.270,\ \chi^2(5) = 9.31(0.10)$$

Panel 5: Imposing weak exogeneity of income with respect to the long-run coefficients

$$ph = \underset{(0.22)}{1.42}d + \underset{(0.83)}{3.79}yh - \underset{(1.51)}{6.48}h$$
$$d = 0.64ph - \underset{(1.67)}{3.58}R + \underset{(0.13)}{0.27}th + \underset{(0.10)}{0.64}h$$
$$\alpha_{1,ph} = \underset{(0.03)}{-0.20},\ \alpha_{1,d} = \underset{(0.02)}{-0.09},\ \alpha_{2,d} = \underset{(0.01)}{-0.05}$$
$$LogL = 986.282,\ \chi^2(7) = 11.28(0.13)$$

The sample is 1986Q2 to 2012Q4, 107 observations

7 Results and Estimation of a Regional Model for Housing Prices

Aggregate models may do a good job in explaining the general trends in regional housing markets, particularly in societies with well developed labor markets, a mobile population, and advanced credit systems. Regional housing markets are then likely to have many common features, e.g. the same type of cointegrating relationships, as the aggregate market. On the other hand, it is unlikely that one and the same econometric model is able to capture well the many nuances that characterize regional markets. To investigate and examine the regional markets, regional data are incorporated into the (updated and extended) A&J dataset in stepwise fashion. First, regional data on housing prices are included at the expense of nationwide housing prices, and the Johansen trace test is conducted to find the cointegrating space. Second, the regional aspect of the model is extended by also including data on regional debt into the aggregated set, and the trace test is once again conducted. This approach is inspired by A&J (2013), which first conduct trace test to fixate the cointegrating relationship before estimating a simultaneous VAR of housing prices and debt. Finally, in section 7.2, an ECM is formulated and estimated, before instrumental variable (IV) estimation is performed in section 7.3.

Another aspect of the analysis is to observe the impact of the latest financial crisis. Norway, though it was not as severly affected as many other countries, was, nevertheless, affected, with real housing prices falling almost 13 percent. In this regard the tests and estimations are performed in two blocks: one ending before the outbreak of the financial crisis, i.e. 2007Q4, and the other containing the full sample, i.e. ending in 2012Q4; effectively including the crisis years. Such a division gives an instructive view of how the housing market reacted in the face of the financial crisis.

7.1 Regional Trace Tests

As explained in section 4.3.1 the Johansen trace test is used to find the dimension of the cointegrating space, i.e. the cointegration rank. As just noted, the trace tests are conducted on the dataset used by A&J (2013), but where national housing prices, and also national debt, are substituted with regional data on housing prices and debt. In this way the model incorporates two important regional trends, and I test the hypothesis that they cointegrate with national explanatory variables. This means that the regional variables on household income and wealth, which was also duly described in chapter 3, are not used in the cointegration analysis.

7.1.1 Regional Housing Prices in the Aggregate Model

A natural first step in the regional analysis is to include regional housing prices in the model. Thus, substituting the national housing price data with regional housing price data from the three regions, while keeping the aggregate data for the remaining variables.

Oslo & Akershus
Table 6 reports the trace test results for the Oslo & Akershus region. The trace test provides rather convincing statistical evidence of two cointegrating relationships for the sample ending in 2007Q4, but the data suggests only one cointegrating relationship when the full sample is used. That the test values become weaker when the full sample is used is not unexpected, in particular when we consider that the financial crisis is contained in the last part of the sample. The test diagnostics reported are satisfying, with the only remark concerning the AR 1-4 test for the sample 1995Q1–2007Q4, for which the null hypothesis of no autocorrelation is rejected at a five percent significance level.

The South-West
The trace test results of the South-West in Table 7, mirrors those of Oslo & Akershus in that two cointegrating relationships are found in the sample ending in 2007Q4, while only one are found for the sample ending in 2012Q4. The standard test diagnostics yield encouraging results, suggesting that the inference based on the trace tests is valid.

Northern Norway
For Northern Norway, see Table 8, the test results marginally indicate rejection of the null hypothesis of two cointegrating relationship in favor of the alternative hypothesis of more than three cointegrating relationships. The general pattern is, nonetheless, similar to that of the two other regions, in that the test values get weaker when estimating over the full sample, i.e. evidence of (only) one cointegrating relationship remains for the sample 1995Q1–2012Q4. The standard test diagnostics give an overall encouraging view of the model, but where the null hypothesis of homoskedasticity is rejected in the extended sample when applying a five percent significance level.

7.1.2 Regional Housing Prices and Debt in the Aggregate Model

The next step is to strengthen the regional aspect by including data for regional debt in the aggregate model. This means that the two most important theoretical variables in the model, housing prices (ph) and debt (d), now contain regional variation. The test procedure is identical to that of the previous section.

Table 6: Trace test for cointegration with regional housing prices in Oslo & Akershus

Estimation period:			1995Q1–2007Q4		1995Q1–2012Q4	
H_0	H_A	5%-critical value	Eigenvalue: λ_i	λ_{trace}	Eigenvalue: λ_i	λ_{trace}
$r = 0$	$r \geq 1$	64.48	0.71	112.77	0.47	75.97
$r \leq 1$	$r \geq 2$	40.95	0.46	48.34	0.22	30.18
$r \leq 2$	$r \geq 3$	20.89	0.27	16.60	0.16	12.28
Diagnostics			Test statistic	Value (p-value)	Test statistic	Value (p-value)
Vector AR 1-4(5) test:[a]			$F(36, 39)$	2.21 (0.01)*	$F(45, 89)$	1.08 (0.37)
Vector Normality test:			$\chi^2(6)$	4.41 (0.62)	$\chi^2(6)$	8.71 (0.19)
Vector Hetero test:			N.A.		$F(270, 133)$	1.14 (0.20)

[a] An AR 1-4 test is conducted for the sample 1995Q1–2007Q4, while an AR 1-5 test is conducted for the sample 1995Q1–2012Q4.

Table 7: Trace test for cointegration with regional housing prices in the South-West

Estimation period:			1995Q1–2007Q4		1995Q1–2012Q4	
H_0	H_A	5%-critical value	Eigenvalue: λ_i	λ_{trace}	Eigenvalue: λ_i	λ_{trace}
$r = 0$	$r \geq 1$	64.48	0.61	103.17	0.48	74.06
$r \leq 1$	$r \geq 2$	40.95	0.53	53.91	0.21	26.96
$r \leq 2$	$r \geq 3$	20.89	0.25	15.09	0.13	10.21
Diagnostics			Test statistic	Value (p-value)	Test statistic	Value (p-value)
Vector AR 1-4(5) test:[a]			$F(36, 39)$	0.94 (0.57)	$F(45, 89)$	0.92 (0.61)
Vector Normality test:			$\chi^2(6)$	8.74 (0.09)	$\chi^2(6)$	8.92 (0.18)
Vector Hetero test:			N.A.		$F(270, 133)$	1.10 (0.26)

[a] An AR 1-4 test is conducted for the sample 1995Q1–2007Q4, while an AR 1-5 test is conducted for the sample 1995Q1–2012Q4.

Table 8: Trace test for cointegration with regional housing prices in Northern Norway

Estimation period:			1995Q1–2007Q4		1995Q1–2012Q4	
H_0	H_A	5%-critical value	Eigenvalue: λ_i	λ_{trace}	Eigenvalue: λ_i	λ_{trace}
$r = 0$	$r \geq 1$	64.48	0.59	104.17	0.42	71.69
$r \leq 1$	$r \geq 2$	40.95	0.50	57.24	0.27	32.47
$r \leq 2$	$r \geq 3$	20.89	0.33	21.05	0.13	9.67
Diagnostics			Test statistic	Value (p-value)	Test statistic	Value (p-value)
Vector AR 1-4(5) test:[a]			$F(36, 39)$	1.72 (0.05)	$F(45, 89)$	1.22 (0.21)
Vector Normality test:			$\chi^2(6)$	8.40 (0.21)	$\chi^2(6)$	6.69 (0.35)
Vector Hetero test:			N.A.		$F(270, 133)$	1.32 (0.04)*

[a] An AR 1-4 test is conducted for the sample 1995Q1–2007Q4, while an AR 1-5 test is conducted for the sample 1995Q1–2012Q4.

Oslo & Akershus
As reported in Table 9, the test values for the first sample show a dramatic fall compared to those observed in section 7.1.1, while the full sample display a remarkable stability in the test values. Thus, the test provides formal evidence of (only) one cointegrating relationship for both samples. Note, however, that the test results are quite similar to those produced in the re-estimation in section 6.1. In that analysis, the Johansen trace test did not provide any statistical evidence of two cointegrating relationships, inference based on the magnitude of the eigenvalues and the corresponding standard errors suggested the existence of two cointegrating relationships. So, even though the trace test may fail to formally support the assumption of two cointegrating relationship, it may nevertheless be supported by further tests and estimations. The test diagnostics indicate that inference based on the trace tests is valid, as long as we apply a one percent significance level when assessing the AR-tests.

The South-West
The same general pattern appears in the South-West, with the test providing formal evidence for one cointegrating relationship. Also for the South-West region the standard test battery yields encouraging results, with the only noteworthy mentioning being the normality test for which it is needed at one percent significance level to keep the null hypothesis of normally distributed disturbances. The results are reported in Table 10.

Northern Norway
As seen in Table 11, Northern Norway differs to some extent from the two other regions. Just as in section 7.1.1 the test suggests that there exists (more than) three cointegrating relationships when investigating the sample 1994Q1–2007Q4. Over the full sample the results are similar to those presented earlier, as only one cointegrating relationship stands out as significant. The reported test diagnostics are satisfying.

Furthermore, the marked fall in test values might suggest that the housing price correction in Northern Norway were more dramatic, and that the housing market in Northern Norway went through a more severe transition in response to the financial crisis. As cointegrating relationships disappear, one might suggest that some kind of equilibrium correction mechanism is lost, and with reference to Abraham and Hendershott (1996), this might indicate a bubble. We return to this issue in the next section.

Table 9: Trace test for cointegration with regional housing prices and debt in Oslo & Akershus

Estimation period:			1995Q1–2007Q4		1995Q1–2012Q4	
H_0	H_A	5%-critical value	Eigenvalue: λ_i	λ_{trace}	Eigenvalue: λ_i	λ_{trace}
$r = 0$	$r \geq 1$	64.48	0.44	61.88	0.38	68.79
$r \leq 1$	$r \geq 2$	40.95	0.36	31.80	0.27	33.84
$r \leq 2$	$r \geq 3$	20.89	0.15	8.74	0.14	11.14
Diagnostics			Test statistic	Value (p-value)	Test statistic	Value (p-value)
Vector AR 1-4(5) test:[a]			$F(36, 39)$	1.88 (0.03)*	$F(45, 89)$	1.86 (0.01)*
Vector Normality test:			$\chi^2(6)$	3.64 (0.72)	$\chi^2(6)$	7.91 (0.24)
Vector Hetero test:			N.A.		$F(270, 133)$	0.90 (0.77)

[a] An AR 1-4 test is conducted for the sample 1995Q1–2007Q4, while an AR 1-5 test is conducted for the sample 1995Q1–2012Q4.

Table 10: Trace test for cointegration with regional housing prices and debt in the South-West

Estimation period:			1995Q1–2007Q4		1995Q1–2012Q4	
H_0	H_A	5%-critical value	Eigenvalue: λ_i	λ_{trace}	Eigenvalue: λ_i	λ_{trace}
$r = 0$	$r \geq 1$	64.48	0.42	61.64	0.35	61.26
$r \leq 1$	$r \geq 2$	40.95	0.31	33.05	0.26	29.91
$r \leq 2$	$r \geq 3$	20.89	0.23	13.83	0.11	8.01
Diagnostics			Test statistic	Value (p-value)	Test statistic	Value (p-value)
Vector AR 1-4(5) test:[a]			$F(36, 39)$	1.46 (0.12)	$F(45, 89)$	1.44 (0.07)
Vector Normality test:			$\chi^2(6)$	15.89 (0.01)*	$\chi^2(6)$	8.54 (0.20)
Vector Hetero test:			N.A.		$F(270, 133)$	0.82 (0.91)

[a] An AR 1-4 test is conducted for the sample 1995Q1–2007Q4, while an AR 1-5 test is conducted for the sample 1995Q1–2012Q4.

Table 11: Trace test for cointegration with regional housing prices and debt in Northern Norway

Estimation period:			1995Q1–2007Q4		1995Q1–2012Q4	
H_0	H_A	5%-critical value	Eigenvalue: λ_i	λ_{trace}	Eigenvalue: λ_i	λ_{trace}
$r = 0$	$r \geq 1$	64.48	0.65	117.93	0.52	83.43
$r \leq 1$	$r \geq 2$	40.95	0.47	62.91	0.29	31.16
$r \leq 2$	$r \geq 3$	20.89	0.44	30.34	0.08	6.22
Diagnostics			Test statistic	Value (p-value)	Test statistic	Value (p-value)
Vector AR 1-4(5) test:[a]			$F(36, 39)$	1.03 (0.46)	$F(45, 89)$	1.20 (0.23)
Vector Normality test:			$\chi^2(6)$	5.01 (0.54)	$\chi^2(6)$	6.01 (0.42)
Vector Hetero test:			N.A.		$F(270, 133)$	1.06 (0.37)

[a] An AR 1-4 test is conducted for the sample 1995Q1–2007Q4, while an AR 1-5 test is conducted for the sample 1995Q1–2012Q4.

7.2 ECM Modelling

Even though the formal tests do not always indicate two cointegrating relationships, the broader analysis indicate that we can continue our analysis under the assumption of two cointegrating relationships. The next step is therefore to investigate if there is any evidence of an error correction mechanism in the housing market with regards to prices. This is done by formulating an ECM, where the ECM term is captured by the parsimonious housing price model in A&J (2013):

$$ECM_t^{ph} = ph_t - 0.98d_t - 1.69yh_t + 3.03h_{t-1} \tag{7.2.1}$$

The justification for using this term is due to its robustness, as the results in A&J (2013) was fairly easily reproduced in section 6.1. While the variable contains information of the stabilizing dynamics of the model, the short-run dynamics of the model is captured by a distributed lag of housing price growth and current debt growth. This is rationalized by the fact that, under the assumption of a cointegrating relationship between housing prices and debt, any factors other than debt, e.g. household income (yh^i), wealth (w^i) and interest rates (R), will affect housing prices through the debt channel. Thus, by including lagged housing price growth, the model hopefully incorporates any effect that other variables besides current debt growth have on housing price growth. Finally, the model contains three variables correcting for any seasonality in the series.

$$\Delta ph_t^i = \phi_0 + \sum_{i=1}^{4} \phi_i \Delta ph_{t-i}^i + \beta_1 \Delta d_t^i + \alpha_1 ECM_t^{ph} + \sum_{j=1}^{3} \gamma_j CSeasonal_j + \varepsilon_t \tag{7.2.2}$$

Due to the reassuring results from the re-estimation of A&J (2013) in section 6.1 and the regional trace tests in the previous section, two separate ECM terms are used:

$$ECM_{t-1}^{i_1} = ph_t^i - 0.98d_t - 1.69yh_t + 3.03h_{t-1} \tag{7.2.3}$$

$$ECM_{t-1}^{i_2} = ph_t^i - 0.98d_t^i - 1.69yh_t + 3.03h_{t-1} \tag{7.2.4}$$

Thus, the ECM-term, $ECM_{t-1}^{i_1}$, in equation (7.2.3), aggregate housing prices are substituted with regional housing prices, while the $ECM_{t-1}^{i_2}$ in equation (7.2.4) is more ambitious by including regional data on housing prices and debt. In the continuation, Model 1 is referred to as the model where the ECM term includes only regional prices, while Model 2 is associated with $ECM^{i_2}_{t-1}$ where both regional prices and debt are incorporated. As before we investigate the model in a pre-crisis sample ending in 2007Q4, and a sample ending in 2012Q3 including the financial crisis.

Oslo & Akershus

The results of the ECM estimation of Oslo & Akershus are found in Table 12. $ECM^{OA_1}_{t-1}$ indicates that a one percentage deviation from (the housing price) equilibrium will result in an equilibrium correction of 7.6 percent in each quarter, indicating that it will take approximately three years until housing prices are back in equilibrium. This adjustment is slower than the corresponding adjustment coefficients found in Jacobsen and Naug (2004*a*) and Anundsen and Jansen (2013*b*) of 12 percent and 24 percent, respectively. It is encouraging to see that the adjustment coefficient stays significant when the full sample is used. Furthermore the short-run dynamics indicate that a percentage increase in the growth of debt will increase housing price growth by 1.3 percent. Once again, the coefficient remains stable and significant when considering the full sample. The standard test diagnostics yields results indicating that inference is valid.

For the model including data on regional housing prices and debt the results are slightly less significant for the parameters of interest, but nevertheless significant. The equilibrium correction coefficient is lower in absolute value than for the previous model, 6.7 percent and 4.8 percent for the two samples respectively, while the growth in debt are almost identical as before. Still, the standard test diagnostics produce inspiring results. Nonetheless, we must accept a significance level of one percent in order to keep the null hypothesis of no autocorrelation when considering the full sample of Model 2.

Lagged price growth, Δph^{OA}_{t-1}, seems to have some explanatory power of the formation of housing prices on the basis of its high t-value. Thus, if housing prices increased by one percent last period, housing prices today would increase by 0.3 percent.

In Figure 8(a) the fitted values and scaled residuals are depicted for the Model 2 version of the ECM. While the model does a good job in explaining actual housing price growth in the first part of the sample, i.e. until 2007Q4, it does a markedly poorer job when including the last five years of the sample. The 2008Q4 observation seems to be an obvious outlier in the sample, as it is associated with the outbreak of the financial crisis, and a task for further research would be to include a dummy variable for this observation to see how the model fares when excluding it.

The South-West

Although the coefficient of the ECM-terms are a bit lower, and a little less significant, in all panels in Table 13, the estimation results for the South-West give the same overall picture as Oslo & Akershus. That the adjustment coefficients drop in absolute value from 5.6 to 4.2, and from 5.4 to 3.9 in the two panels are as expected, taken into account the outbreak of the financial crisis and the subsequent turmoil in financial markets which i.a. led to a tightening in the credit provision. The debt growth once again proves very significant, and stable over both samples and in both models. Also for the housing price formation in the South-West lagged housing price growth seems to be a driver, with a percentage increase in lagged housing price growth to increase housing price growth this quarter by 0.3 percent. The standard test apparatus suggests inference is valid.

The fitted values and scaled residuals for ECM for the South-West are given in Figure 8(b). It seems like the model does a marginally poorer job in explaining actual housing price growth than the model for Oslo & Akershus, but it shows the same overall trends, as the model falls apart when including the financial crisis period.

Northern Norway

In Northern Norway, see Table 14, the adjustment coefficient in Model 1 is higher than in the two preceding regions, with housing prices expected to adjust by 8.4 percent each quarter of a percentage deviation from a housing price equilibrium. This adjustment coefficient increases to 9.7 percent when considering the full sample. Thus, it will take approximately ten quarters before equilibrium is restored. When considering Model 2, the adjustment coefficient turns out lower: 5.9 percent and 4.7 percent in the two samples respectively. However, they are still significantly different from zero. From the trace test results in the previous section, it was inferred that due to the dramatic change in test values in Northern Norway it might be that the housing market was more affected than the markets in the two other regions. This may still be the case, but in relation to a bubble in the housing market in Northern Norway, the error correction coefficients in Model 1 suggest a relatively fast recovery of disequilibrium housing prices, indicating a relationship between housing prices and fundamentals, and thus, in relation to Abraham and Hendershott's notion of a "bubble burster", no bubble. The effect of debt growth on housing price growth is similar to the other regions, and is significant in all estimations. Once again, the standard tests produce encouraging results.

The fitted values and scaled residuals for the ECM for Northern Norway are shown in Figure 8(c). Once again the figures demonstrate the same features, as the model does a fairly good job in explaining housing price growth over the first part of the sample, but a poorer job over the years that were marked by the financial crisis.

Table 12: An ECM for housing prices in Oslo & Akershus

	Model 1[a]				Model 2[b]			
	1994Q1–2007Q4		1994Q1–2012Q4		1994Q1–2007Q4		1994Q1–2012Q4	
Variable	Coefficient	t-value	Coefficient	t-value	Coefficient	t-value	Coefficient	t-value
$Constant$	0.943	3.95	0.906	4.21	0.898	3.78	0.655	3.36
Δph_{t-1}^{OA}	0.262	2.23	0.232	2.52	0.277	2.34	0.242	2.52
Δph_{t-2}^{OA}	0.117	0.97	0.024	0.25	0.127	1.03	0.012	0.12
Δph_{t-3}^{OA}	−0.120	−1.03	−0.057	−0.60	−0.114	−0.97	−0.061	−0.62
Δph_{t-4}^{OA}	0.205	1.88	0.119	1.31	0.218	1.98	0.121	1.28
$ECM_{t-1}^{OA_1}$	−0.076	−3.95	−0.073	−4.22	–	–	–	–
$ECM_{t-1}^{OA_2}$	–	–	–	–	−0.067	−3.78	−0.048	−3.37
Δd_t^{OA}	1.251	5.07	1.269	6.62	1.252	4.96	1.148	5.97
$CSeasonal_1$	0.012	0.97	0.023	2.38	0.012	0.94	0.025	2.54
$CSeasonal_2$	−0.011	−0.80	−0.002	−0.24	−0.012	−0.89	−0.000	−0.01
$CSeasonal_3$	−0.019	−1.56	−0.005	−0.56	−0.019	−1.59	−0.006	−0.64
Log-likelihood	152.82		203.41		152.21		200.39	
$F(9,46)$	20.15 (0.00)**		–		19.61 (0.00)**		–	
$F(9,66)$	–		22.10 (0.00)**		–		19.86 (0.00)**	
Diagnostics	Test statistic	Value (p-value)	Test statistic	Value (p-value)	Test statistic	Value (p-value)	Test statistic	Value (p-value)
AR 1-4(5) test:[c]	$F(4,42)$	1.22 (0.31)	$F(5,61)$	1.57 (0.18)	$F(4,42)$	1.01 (0.41)	$F(5,61)$	2.48 (0.04)*
Normality test:	$\chi^2(2)$	0.43 (0.81)	$\chi^2(2)$	2.36 (0.31)	$\chi^2(2)$	0.43 (0.81)	$\chi^2(2)$	2.19 (0.33)
Hetero test:	$F(15,40)$	0.64 (0.82)	$F(15,60)$	0.91 (0.56)	$F(15,40)$	0.67 (0.80)	$F(15,60)$	0.93 (0.54)
Estimation method	OLS							

[a] ECM with the accompanying ECM term $ECM_{t-1}^{OA_1} = ph^{OA} - 0.98d - 1.69yh + 3.03h$; i.e. incl. regional housing prices.
[b] ECM with the accompanying ECM term $ECM_{t-1}^{OA_2} = ph^{OA} - 0.98d^{OA} - 1.69yh + 3.03h$; incl. regional housing prices and debt.
[c] For the sample 1994Q1–2007Q4 an AR 1-4 test is conducted, while an AR 1-5 test is conducted for the extended sample ending in 2012Q4.

Table 13: An ECM for housing prices in the South-West

	Model 1[a]				Model 2[b]			
	1994Q1–2007Q4		1994Q1–2012Q4		1994Q1–2007Q4		1994Q1–2012Q4	
Variable	Coefficient	t-value	Coefficient	t-value	Coefficient	t-value	Coefficient	t-value
$Constant$	0.680	2.07	0.504	2.28	0.716	2.17	0.529	2.35
Δph_{t-1}^{SW}	0.348	2.63	0.309	2.99	0.342	2.60	0.305	2.96
Δph_{t-2}^{SW}	0.130	0.98	0.021	0.19	0.129	0.98	0.021	0.19
Δph_{t-3}^{SW}	−0.276	−2.11	−0.157	−1.45	−0.280	−2.15	−0.160	−1.48
Δph_{t-4}^{SW}	0.123	0.96	0.055	0.53	0.118	0.92	0.051	0.49
$ECM_{t-1}^{SW_1}$	−0.056	−2.05	−0.042	−2.29	–	–	–	–
$ECM_{t-1}^{SW_2}$	–	–	–	–	−0.054	−2.15	−0.039	−2.36
Δd_t^{SW}	1.166	3.64	1.195	5.01	1.191	3.71	1.207	5.05
$CSeasonal_1$	0.017	0.96	0.031	2.51	0.017	0.97	0.032	2.54
$CSeasonal_2$	−0.029	−1.58	−0.007	−0.52	−0.029	−1.59	−0.007	−0.53
$CSeasonal_3$	−0.037	−2.38	−0.016	−1.29	−0.037	−2.39	−0.016	−1.30
Log-likelihood	138.63		184.73		138.86		184.91	
$F(9,46)$	15.28 (0.00)**		–		15.28 (0.00)**		–	
$F(9,66)$	–		16.63 (0.00)**		–		16.74 (0.00)**	
Diagnostics	Test statistic	Value (p-value)	Test statistic	Value (p-value)	Test statistic	Value (p-value)	Test statistic	Value (p-value)
AR 1-4(5) test[c]	$F(4,42)$	1.35 (0.27)	$F(5,61)$	1.93 (0.10)	$F(4,42)$	1.40 (0.25)	$F(5,61)$	1.96 (0.10)
Normality test:	$\chi^2(2)$	1.73 (0.42)	$\chi^2(2)$	3.90 (0.14)	$\chi^2(2)$	1.64 (0.44)	$\chi^2(2)$	3.90 (0.14)
Hetero test:	$F(15,40)$	0.56 (0.89)	$F(15,60)$	0.63 (0.84)	$F(15,40)$	0.57 (0.88)	$F(15,60)$	0.64 (0.83)
Estimation method	OLS							

[a] ECM with the accompanying ECM term $ECM_{t-1}^{SW_1} = ph^{SW} - 0.98d - 1.69yh + 3.03h$; i.e. incl. regional housing prices.

[b] ECM with the accompanying ECM term $ECM_{t-1}^{SW_2} = ph^{SW} - 0.98d^{SW} - 1.69yh + 3.03h$; incl. regional housing prices and debt.

[c] For the sample 1994Q1–2007Q4 an AR 1-4 test is conducted, while an AR 1-5 test is conducted for the extended sample ending in 2012Q4.

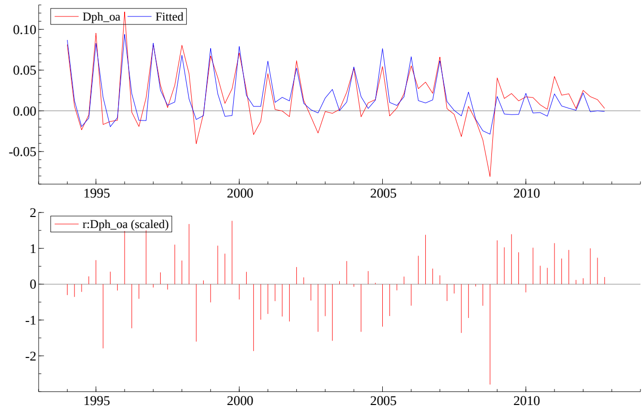
Table 14: An ECM for housing prices in Northern Norway

	Model 1[a]				Model 2[b]			
	1994Q1–2007Q4		1994Q1–2012Q4		1994Q1–2007Q4		1994Q1–2012Q4	
Variable	Coefficient	t-value	Coefficient	t-value	Coefficient	t-value	Coefficient	t-value
$Constant$	1.011	1.94	1.160	2.50	0.792	2.08	0.629	2.13
Δph_{t-1}^{NN}	0.402	2.99	0.394	3.66	0.397	2.98	0.379	3.48
Δph_{t-2}^{NN}	−0.024	−0.17	−0.079	−0.67	−0.023	−0.16	−0.096	−0.80
Δph_{t-3}^{NN}	−0.137	−0.97	−0.043	−0.37	−0.150	−1.07	−0.060	−0.51
Δph_{t-4}^{NN}	0.131	1.01	0.097	0.92	0.127	0.99	0.064	0.61
$ECM_{t-1}^{NN_1}$	−0.084	−1.94	−0.097	−2.51	–	–	–	–
$ECM_{t-1}^{NN_2}$	–	–	–	–	−0.059	−2.08	−0.047	−2.14
Δd_t^{NN}	1.287	3.77	1.330	4.72	1.424	3.90	1.314	4.58
$CSeasonal_1$	0.022	1.08	0.036	2.45	0.018	0.91	0.039	2.62
$CSeasonal_2$	−0.025	−1.33	−0.006	−0.41	−0.026	−1.37	−0.008	−0.51
$CSeasonal_3$	−0.020	−1.09	−0.002	−0.14	−0.022	−1.22	−0.005	−0.35
Log-likelihood	130.65		175.81		130.96		174.91	
$F(9,46)$	14.03 (0.00)**		–		14.24 (0.00)**		–	
$F(9,66)$	–		16.90 (0.00)**		–		16.33 (0.00)**	
Diagnostics	Test statistic	Value (p-value)	Test statistic	Value (p-value)	Test statistic	Value (p-value)	Test statistic	Value (p-value)
AR 1-4(5) test:[c]	$F(4,42)$	0.70 (0.60)	$F(5,61)$	0.82 (0.54)	$F(4,42)$	0.83 (0.51)	$F(5,61)$	0.87 (0.41)
Normality test:	$\chi^2(2)$	2.75 (0.25)	$\chi^2(2)$	3.34 (0.19)	$\chi^2(2)$	1.25 (0.54)	$\chi^2(2)$	3.12 (0.21)
Hetero test:	$F(15,40)$	1.20 (0.31)	$F(15,60)$	1.79 (0.06)	$F(15,40)$	1.37 (0.21)	$F(15,60)$	2.00 (0.03)*
Estimation method	OLS							

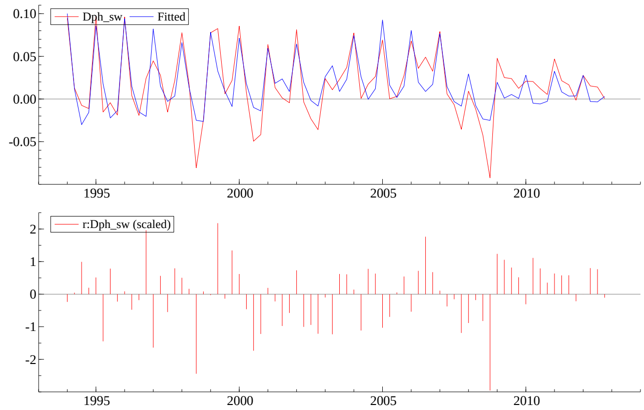
[a] ECM with the accompanying ECM term $ECM_{t-1}^{NN_1} = ph^{NN} - 0.98d - 1.69yh + 3.03h$; i.e. incl. regional housing prices.

[b] ECM with the accompanying ECM term $ECM_{t-1}^{NN_2} = ph^{NN} - 0.98d^{NN} - 1.69yh + 3.03h$; incl. regional housing prices and debt.

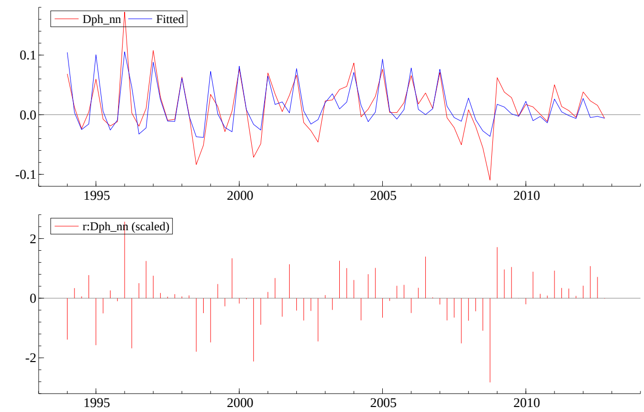
[c] For the sample 1994Q1–2007Q4 an AR 1-4 test is conducted, while an AR 1-5 test is conducted for the extended sample ending in 2012Q4.



(a) Fitted values and scaled residuals for the ECM model for Oslo & Akershus



(b) Fitted values and scaled residuals for the ECM model for South-West



(c) Fitted values and scaled residuals for the ECM model for Northern Norway

Figure 8: Fitted values and scaled residuals for the three regional ECM models including regional housing prices and debt (i.e. Model 2) over the period 1994Q1–2012Q4.

7.3 Instrumental Variable Estimation

So far, our models have been conditional on the within quarter (contemporaneous) growth in credit. This is not strictly consistent with our theoretical framework, and is also contradicted by the simultaneous equations model presented by A&J, which was reviewed earlier. In this section, I assume that other factors excluded from the ECM formulation in equation (7.2.2) may affect housing prices by being instruments for the debt variable, Δd_t^i. As noted, Anundsen and Jansen (2013*b*) find evidence for self-reinforcing effects between housing prices and debt, and model them in a simultaneous system. Also, Δd_t^i is included as an endogenous variable in the model, and instrumental variable estimation is executed in order to try explain the drivers for household debt. This would give further information of what is the main determinants of housing prices.

Instrumental variable estimation is based on relevant and valid instruments:

i. *Relevance*

An instrument is relevant if it is (highly) correlated with the variable it acts as an instrument for. The higher correlation coefficient, the stronger the instrument.

ii. *Validity*

An instrument is valid if it is uncorrelated with the error term in the equation. If the validity condition does not hold, IV estimation will not be consistent, i.e. the IV estimator will not converge in probability to the true value.

In the IV-estimation of the ECM equations the following variables are used as instruments: housing price growth (Δyh_t^i), lagged debt growth (Δd_{t-1}^i), lagged debt (d_{t-1}^i), wealth (w_t^i), lagged wealth (w_{t-1}^i), the real interest rate (R_t), and the the real interest rate in the previous period (R_{t-1}). While most of the instruments are likely to satisfy the relevance condition, there might be more uncertainty with regards to the validity condition. The specification test reported in the tables 15–17 below is the Sargan test. The null hypothesis is that the instruments are valid. Hence, a significant Sargan test indicates that the instruments are valid. If the Sargan test comes out insignificant this might suggest that some of the instruments are invalid, but as emphasized by Bårdsen and Nymoen (2011, pg. 231–232), another interpretation is that one or more instruments might have explanatory power of the residuals because they are omitted variables in the structural equation. In the last case, a re-specification of the structural equations might be advisable.

Oslo & Akershus

It is clear that the specification test reported in Table 15 indicates valid instruments in both models, and over both sample lengths. Even though there is a sharp decline in the reported p-values when considering the full sample, the tests are still significant. All the standard tests are also significant, and valid inference is justified.

When endogenizing debt growth with the help of the instruments, 2SLS estimation shows that the estimated effect of a percentage change in growth of debt increases somewhat compared to the estimation results obtained in the previous section. t-values remain high. However, in the model including both regional housing prices and debt there is a marked fall in the coefficient from 1.6 percent for the sample ending in 2007Q4 to 1.1 percent for the sample ending in 2012Q3. This might reflect the fact that credit provision were tightened after the outbreak of the financial crisis in 2008, hence, reducing the pressure on housing prices. A similar decrease is seen in Model 1, but not to the same extent.

The adjustment coefficients are recognizable from previous estimations. That the degree of equilibrium correction decreases when considering the full sample seems reasonable as the housing market suffered somewhat and credit tightened in the aftermath of the crisis.

The South-West

The same overall pattern as in Oslo & Akershus is seen in the South-West with significant specification tests, test diagnostics, and the trends concerning the debt and ECM variables, see Table 16. However, the relatively sharp drop in the debt coefficient in Model 2 is not matched in the South-West region; i.e. the coefficient falls from 1.2 to 1.1. This might indicate that the credit did not dry up in the South-West in the face of the crisis, for instance because of high economic activity in the region, and pressure in the housing market due to net positive migration. The ECM coefficients fall about two percentage points in both Model 1 and 2 when extending the sample period, from 6.5 percent to 4.1 percent and 5.9 percent to 3.8 percent, respectively.

Northern Norway

Northern Norway exhibit some differences compared to the other two regions, see Table 17. For instance, while the debt effect increases relative to the OLS estimation results, it does increase when extending the data sample in Model 1. This is in sharp contrast to Oslo & Akershus and the South-West. Interestingly, the estimation results suggest that debt financing has more to say for housing prices in Northern Norway than in the two other regions. However, in Model 2, the coefficient falls, just as in the other two regions, indicating that the difference might not be significant, and that one should not read too much in to it. Another interesting aspect is the ECM coefficient in Model 1, which increases from 10.1 to 11.8, indicating a faster equilibrium correction when estimating over the full sample. Along the lines proposed here, this change is counterintuitive, but it may, nevertheless, indicate regional differences. However, once again, in Model 2, the same decrease in the adjustment coefficient is seen when comparing the two samples. The test diagnostics imply no model mis-specification.

Table 15: Instrumental variable (IV) estimation of household debt in the ECM for housing prices in Oslo & Akershus[a]

	Model 1[b]				Model 2[c]			
	1994Q2–2007Q4		1994Q2–2012Q3		1994Q2–2007Q4		1994Q2–2012Q3	
Variable	Coefficient	t-value	Coefficient	t-value	Coefficient	t-value	Coefficient	t-value
$\Delta d_t^{OA}(Y)$[d]	1.531	4.68	1.305	5.15	1.588	4.66	1.104	4.48
$Constant$	1.148	3.93	0.971	3.86	1.138	3.84	0.662	2.98
Δph_{t-1}^{OA}	0.276	2.31	0.236	2.51	0.297	2.45	0.246	2.51
Δph_{t-2}^{OA}	0.157	1.24	0.026	0.26	0.176	1.36	0.005	0.05
Δph_{t-3}^{OA}	−0.091	−0.75	−0.045	−0.46	−0.080	−0.65	−0.054	−0.52
Δph_{t-4}^{OA}	0.164	1.37	0.097	0.99	0.176	1.46	0.117	1.13
$ECM_{t-1}^{OA_1}$	−0.093	−3.93	−0.078	−3.86	–	–	–	–
$ECM_{t-1}^{OA_2}$	–	–	–	–	−0.085	−3.83	−0.049	−2.98
$CSeasonal_1$	0.011	0.80	0.025	2.44	0.010	0.73	0.027	2.55
$CSeasonal_2$	−0.007	−0.52	0.004	0.34	−0.008	−0.60	0.000	0.00
$CSeasonal_3$	−0.019	−1.54	−0.004	−0.44	−0.020	−1.57	−0.005	−0.52
Diagnostics	Test statistic	Value (p-value)	Test statistic	Value (p-value)	Test statistic	Value (p-value)	Test statistic	Value (p-value)
AR 1-4(5) test:[e]	$F(4,41)$	1.35 (0.27)	$F(5,59)$	1.95 (0.10)	$F(4,41)$	1.10 (0.37)	$F(5,59)$	2.74 (0.03)*
Normality test:	$\chi^2(2)$	0.56 (0.76)	$\chi^2(2)$	2.38 (0.30)	$\chi^2(2)$	0.45 (0.80)	$\chi^2(2)$	2.25 (0.32)
Hetero test:	$F(15,39)$	0.76 (0.71)	$F(15,58)$	1.08 (0.40)	$F(15,39)$	0.78 (0.70)	$F(15,58)$	1.00 (0.46)
Specification test:	$\chi^2(6)$	1.41 (0.97)	$\chi^2(6)$	7.94 (0.24)	$\chi^2(6)$	1.59 (0.95)	$\chi^2(6)$	12.29 (0.06)
Estimation method	2SLS							

[a] The instruments for Δd_t^i are: Δyh_t^i, Δd_{t-1}^i, d_{t-1}^i, w_t^i, w_{t-1}^i, R_t, R_{t-1}.
[b] ECM with the accompanying ECM term $ECM_{t-1}^{OA_1} = ph^{OA} - 0.98d - 1.69yh + 3.03h$; i.e. incl. regional housing prices.
[c] ECM with the accompanying ECM term $ECM_{t-1}^{OA_2} = ph^{OA} - 0.98d^{OA} - 1.69yh + 3.03h$; incl. regional housing prices and debt.
[d] The Y denotes that Δd_t^{OA} is incl. as an endogenous variable.
[e] For the sample 1994Q2–2007Q4 an AR 1-4 test is conducted, while an AR 1-5 test is conducted for the extended sample ending in 2012Q3.

Table 16: Instrumental variable (IV) estimation of household debt in the ECM for housing prices in the South-West[a]

	Model 1[b]				Model 2[c]			
	1994Q2–2007Q4		1994Q2–2012Q3		1994Q2–2007Q4		1994Q2–2012Q3	
Variable	Coefficient	t-value	Coefficient	t-value	Coefficient	t-value	Coefficient	t-value
$\Delta d_t^{SW}(Y)$[d]	1.246	2.95	1.142	3.84	1.232	2.96	1.114	3.77
$Constant$	0.780	1.88	0.496	2.02	0.787	1.92	0.504	2.04
Δph_{t-1}^{SW}	0.364	2.64	0.316	2.97	0.355	2.61	0.315	2.96
Δph_{t-2}^{SW}	0.145	1.04	0.018	0.16	0.140	1.02	0.015	0.14
Δph_{t-3}^{SW}	−0.262	−1.94	−0.156	−1.39	−0.269	−2.01	−0.159	−1.42
Δph_{t-4}^{SW}	0.122	0.94	0.058	0.55	0.115	0.90	0.058	0.54
$ECM_{t-1}^{SW_1}$	−0.065	−1.86	−0.041	−2.02	–	–	–	–
$ECM_{t-1}^{SW_2}$	–	–	–	–	−0.059	−1.90	−0.038	−2.03
$CSeasonal_1$	0.017	0.90	0.032	2.45	0.018	0.97	0.033	2.51
$CSeasonal_2$	−0.028	−1.53	−0.008	−0.56	−0.028	−1.55	−0.008	−0.61
$CSeasonal_3$	−0.037	−2.34	−0.016	−1.24	−0.037	−2.36	−0.016	−1.25
Diagnostics	Test statistic	Value (p-value)	Test statistic	Value (p-value)	Test statistic	Value (p-value)	Test statistic	Value (p-value)
AR 1-4(5) test:[e]	$F(4,41)$	2.00 (0.11)	$F(5,59)$	2.15 (0.07)	$F(4,41)$	2.06 (0.10)	$F(5,59)$	2.17 (0.07)
Normality test:	$\chi^2(2)$	1.70 (0.43)	$\chi^2(2)$	3.78 (0.15)	$\chi^2(2)$	1.70 (0.43)	$\chi^2(2)$	3.95 (0.14)
Hetero test:	$F(15,39)$	0.52 (0.91)	$F(15,58)$	0.64 (0.83)	$F(15,39)$	0.53 (0.90)	$F(15,58)$	0.65 (0.82)
Specification test:	$\chi^2(6)$	5.06 (0.54)	$\chi^2(6)$	9.37 (0.15)	$\chi^2(6)$	5.11 (0.53)	$\chi^2(6)$	10.15 (0.12)
Estimation method	2SLS							

[a] The instruments for Δd_t^i are: Δyh_t^i, Δd_{t-1}^i, d_{t-1}^i, w_t^i, w_{t-1}^i, R_t, R_{t-1}.

[b] ECM with the accompanying ECM term $ECM_{t-1}^{SW_1} = ph^{SW} - 0.98d - 1.69yh + 3.03h$; i.e. incl. regional housing prices.

[c] ECM with the accompanying ECM term $ECM_{t-1}^{SW_2} = ph^{SW} - 0.98d^{SW} - 1.69yh + 3.03h$; incl. regional housing prices and debt.

[d] The Y denotes that Δd_t^{SW} is incl. as an endogenous variable.

[e] For the sample 1994Q2–2007Q4 an AR 1-4 test is conducted, while an AR 1-5 test is conducted for the extended sample ending in 2012Q3.

Table 17: Instrumental variable (IV) estimation of household debt in the ECM for housing prices in Northern Norway[a]

	Model 1[b]				Model 2[c]			
	1994Q2–2007Q4		1994Q2–2012Q3		1994Q2–2007Q4		1994Q2–2012Q3	
Variable	Coefficient	t-value	Coefficient	t-value	Coefficient	t-value	Coefficient	t-value
$\Delta d_t^{NN}(Y)$[d]	1.256	3.39	1.404	4.21	1.370	3.33	1.194	3.50
$Constant$	1.214	2.17	1.415	2.82	0.892	2.14	0.682	2.12
Δph_{t-1}^{NN}	0.415	3.12	0.405	3.78	0.408	3.10	0.399	3.63
Δph_{t-2}^{NN}	0.013	0.09	−0.052	−0.44	0.010	0.07	−0.089	−0.73
Δph_{t-3}^{NN}	−0.106	−0.75	−0.023	−0.19	−0.121	−0.87	−0.039	−0.32
Δph_{t-4}^{NN}	0.099	0.76	0.075	0.72	0.089	0.70	0.044	0.42
$ECM_{t-1}^{NN_1}$	−0.101	−2.16	−0.118	−2.82	–	–	–	–
$ECM_{t-1}^{NN_2}$	–	–	–	–	−0.067	−2.14	−0.051	−2.12
$CSeasonal_1$	0.030	1.49	0.040	2.66	0.028	1.36	0.046	2.97
$CSeasonal_2$	−0.023	−1.22	−0.003	−0.23	−0.024	−1.29	−0.008	−0.50
$CSeasonal_3$	−0.020	−1.09	−0.001	−0.07	−0.023	−1.26	−0.004	−0.28
Diagnostics	Test statistic	Value (p-value)	Test statistic	Value (p-value)	Test statistic	Value (p-value)	Test statistic	Value (p-value)
AR 1-4(5) test[e]	$F(4,41)$	0.69 (0.60)	$F(5,59)$	0.64 (0.67)	$F(4,41)$	1.23 (0.31)	$F(5,59)$	0.76 (0.58)
Normality test:	$\chi^2(2)$	2.48 (0.29)	$\chi^2(2)$	2.74 (0.25)	$\chi^2(2)$	1.49 (0.47)	$\chi^2(2)$	3.50 (0.17)
Hetero test:	$F(15,39)$	1.14 (0.36)	$F(15,58)$	1.56 (0.11)	$F(15,39)$	1.35 (0.22)	$F(15,58)$	1.91 (0.04)*
Specification test:	$\chi^2(6)$	6.07 (0.42)	$\chi^2(6)$	4.59 (0.60)	$\chi^2(6)$	5.45 (0.49)	$\chi^2(6)$	6.51 (0.37)
Estimation method	2SLS							

[a] The instruments for Δd_t^i are: Δyh_t^i, Δd_{t-1}^i, d_{t-1}^i, w_t^i, w_{t-1}^i, R_t, R_{t-1}.

[b] ECM with the accompanying ECM term $ECM_{t-1}^{NN_1} = ph^{NN} - 0.98d - 1.69yh + 3.03h$; i.e. incl. regional housing prices.

[c] ECM with the accompanying ECM term $ECM_{t-1}^{NN_2} = ph^{NN} - 0.98d^{NN} - 1.69yh + 3.03h$; incl. regional housing prices and debt.

[d] The Y denotes that Δd_t^{NN} is incl. as an endogenous variable.

[e] For the sample 1994Q2–2007Q4 an AR 1-4 test is conducted, while an AR 1-5 test is conducted for the extended sample ending in 2012Q3.

8 Conclusion

The main objective of this thesis has been to investigate the housing markets in different regions in Norway, to see which factors are the main determinants of housing price formation, and to see if the housing price increase in recent years can be justified by fundamentals. My econometric model builds on Anundsen and Jansen (2013*b*) which establish a housing price equation consisting of debt, income, interest rates, and the housing stock. These results were shown to be robust for an extended and revised dataset, indicating that the housing price model in A&J (2013) gives a good description of the formation of housing prices.

The regional characteristics are incorporated into the aggregated model proposed by A&J (2013) by substituting aggregated housing prices and debt with their regional counterparts. The trace tests suggest that there exist two cointegrating relationships in the different regions for the sample 1995Q1–2007Q4, but the data provide weaker evidence of two cointegrating relationship when considering a sample including the outbreak and the wake of the financial crisis. The test statistics are altogether weaker when including both regional housing prices and debt in the model, and the test indicates proof of only one cointegrating relationship in Oslo & Akershus and the South-West, but as was emphasized in the re-estimation section, further testing may provide evidence pointing towards two cointegrating relationships. In Northern Norway, however, three cointegrating relationships are found in the sample ending in 2007Q4, but only one relationship remains when considering the full sample. This might mean that the turbulent years in the aftermath of the crisis affected the housing market in Northern Norway more than it did in the the two other regions.

In the ECM formulation the corresponding estimation results gave the overall same pattern in all three regions, with some minor deviations when extending the sample length and incorporating more regional variables. The main finding is that housing markets in the different regions are remarkably similar and synchronized. The adjustment coefficients are found to be consistently lower than in earlier findings, cf. Jacobsen and Naug (2004*a*), and they fall somewhat when incorporating the financial crisis in the estimation period. Nevertheless, the coefficients are significantly different from zero according to the data, indicating that a deviation from a housing market equilibrium will produce equilibrium correction. In accordance with Abraham and Hendershott (1996) this implies that housing prices will adjust themselves to stay in touch with fundamentals, yet the correction will be slow. Thus, there is no formal evidence of a bubble in the regional housing markets in Norway – recall that no equilibrium correction would indicate that housing prices would drift independently of fundamentals, which would indicate a bubble according to Stiglitz's (1990) definition – in other words it seems like housing price growth in Norway is driven by fundamentals.

Finally, the IV estimation suggest that other important factors in determining housing prices are income growth, lagged debt growth, lagged debt, wealth, and interest rates. Furthermore, there seems to be reasonable to estimate housing prices and debt in a simultaneous equations system, as there are obviously linkages between the two. Note that the estimation results of the regional ECM models for housing prices in all important respects are robust to the choice of estimation method (OLS and IV/2SLS).

Hence, it seems that the aggregate housing price model is well suited for also capturing and explaining regional development trends. Nonetheless, the regional model indicates that there are contributions to be made from including variables capturing regional heterogeneity in our housing price models. In that way, the results may represent a starting point for more econometric modelling of regional housing markets. For example, extending the analysis by looking at additional regions, a natural extension would be to incorporate more regional variables in the information set. Such an extension would give a richer and more detailed description of the regional dynamics. Another interesting line of attack is examining ripple effects, e.g. how well does regional housing prices in Oslo & Akershus explain housing price movements in Northern Norway, in addition to analyzing the effects of demographic trends.

References

Abraham, J. M. and Hendershott, P. H. (1996), 'Bubbles in Metropolitan Housing Markets', *Journal of Housing Research* **7**(2), 191–207.

Anundsen, A. K. (2010), Boligpriser, forventninger og gjeld – En økonometrisk analyse av samspillet mellom bolig- og kredittmarkedet i Norge, Master's thesis, University of Oslo.

Anundsen, A. K. (2013), 'Econometric Regime Shifts and the US Subprime Bubble', *Journal of Applied Econometrics* . Forthcoming.

Anundsen, A. K. (2014), Housing Markets and Financial Stability, PhD thesis, University of Oslo.

Anundsen, A. K. and Jansen, E. S. (2013*a*), 'Self-Reinforcing Effects Between Housing Prices and Credit', *Journal of Housing Economics* **22**(3), 192–212.

Anundsen, A. K. and Jansen, E. S. (2013*b*), 'Self-Reinforcing Effects Between Housing Prices and Credit: An Extended Version', *Discussion Papers* (756). Oslo: Statistics Norway.

Bårdsen, G. and Nymoen, R. (2011), *Innføring i økonometri*, Fagbokforlaget, Bergen.

Bårdsen, G. and Nymoen, R. (2014), *Videregående emner i økonometri*, Fagbokforlaget, Bergen.

Bernanke, B., Gertler, M. and Gilchrist, S. (1999), The Financial Accelerator in a Quantitative Business Cycle Framework, *in* J. B. Taylor and M. Woodford, eds, 'Handbook of Macroeconomics', Vol. 1, North-Holland, Amsterdam, pp. 1341–1393.

Borgersen, T.-A. and Hungnes, H. (2009), 'Selvforsterkende effekter i bolig- og kredittmarkeder', *Norsk Økonomisk Tidsskrift* (123), 18–33.

Boug, P. and Dyvi, Y. (2008), 'MODAG – En makroøkonomisk modell for norsk økonomi', *Sosiale og økonomiske studier* (111). Oslo: Statistics Norway.

Davidson, R. and MacKinnon, J. G. (2009), *Econometric Theory and Methods*, Oxford University Press, New York.

Dickey, D. A. and Fuller, W. A. (1979), 'Distribution of the Estimators for Autoregressive Time Series with a Unit Root', *Journal of the American Statistical Association* **74**(366a), 427–431.

Doornik, J. A. (2003), 'Asymptotic Tables for Cointegration Tests Based on The Gamma-Distribution Approximation', *Unpublished note* . Nuffield College, Oxford.

Doornik, J. A. and Hendry, D. F. (2009), *Empirical Econometric Modelling – PcGive 13*, Vol. 1, Timberlake Consultants.

Duca, J. V., Muellbauer, J. and Murphy, A. (2011), 'House Prices and Credit Constraints: Making Sense of the US Experience', *The Economic Journal* **121**(552), 533–551.

Eitrheim, Ø. (1993), 'En dynamisk modell for boligprisen i RIMINI', *Penger og Kreditt* **4**, 288–297.

Eitrheim, Ø. and Erlandsen, S. K. (2004), House Price Indices for Norway 1819–2003, *in* Ø. Eitrheim, J. T. Klovland and Q. J. F, eds, 'Historical Monetary Statistics for Norway', number 35 *in* 'Occasional Papers', Norges Bank, chapter 9, pp. 349–378.

Engle, R. F. and Granger, C. W. (1987), 'Co-Integration and Error Correction: Representation, Estimation, and Testing', *Econometrica: Journal of the Econometric Society* **55**(2), 251–276.

Flood, R. P. and Hodrick, R. J. (1990), 'On Testing for Speculative Bubbles', *Journal of Economic Perspectives* **4**(2), 85–101.

Gallin, J. (2006), 'The Long-Run Relationship Between House Prices and Income: Evidence from Local Housing Markets', *Real Estate Economics* **34**(3), 417–438.

Gerdrup, K. R. (2003), Three Episodes of Financial Fragility in Norway Since the 1890s, BIS Working Papers 142, Bank for International Settlements, Basel. Basel.

Girouard, N., Kennedy, M., Van Den Noord, P. and André, C. (2006), 'Recent House Price Developments: The Role of Fundamentals', *OECD Economics Department Working Papers* (475).

Granger, C. W. (1990), *Modelling Economic Series: Readings in Econometric Methodology*, Clarendon Press, Oxford.

Granger, C. W. and Newbold, P. (1974), 'Spurious Regressions in Econometrics', *Journal of Econometrics* **2**(2), 111–120.

Hamilton, J. D. (1994), *Time Series Analysis*, Princeton University Press, Princeton, New Jersey.

Hendry, D. F. (1984), *Econometrics and Quantitative Economics*, Blackwell, Oxford, chapter 8: Econometric Modelling of House Prices in the United Kingdom, pp. 211–252.

Hill, R. C., Griffiths, W. E. and Lim, G. C. (2008), *Principles of Econometrics*, third edn, Wiley & Sons, Hoboken, New Jersey.

IMF (2013), Norway: Selected issues, Country Report 273, Washington D.C. Authors: Cheptea, C. and Shirono, K. and Mordonu, A. and Dowling, T.

Jacobsen, D. H. and Naug, B. E. (2004*a*), 'Hva driver boligprisene?', *Penger og Kreditt* (4), 229–240.

Jacobsen, D. H. and Naug, B. E. (2004*b*), 'Hva påvirker gjeldsveksten i husholdningene?', *Penger og Kreditt* (2), 91–98.

Kiyotaki, N. and Moore, J. (1997), 'Credit Cycles', *Journal of Political Economy* **105**(2), 211–248.

Larsen, E. R. and Sommervoll, D. E. (2004), 'Hva bestemmer boligprisene?', *Samfunnsspeilet* **18**(2), 10–17.

Malpezzi, S. (1999), 'A Simple Error Correction Model of House Prices', *Journal of Housing Economics* **8**(1), 27–62.

Meen, G. (2002), 'The Time-Series Behavior of House Prices: A Transatlantic Divide?', *Journal of Housing Economics* **11**(1), 1–23.

Meen, G. and Andrew, M. (1998), 'On the Aggregate Housing Market Implications of Labour Market Change', *Scottish Journal of Political Economy* **45**(4), 393–419.

Mikhed, V. and Zemčík, P. (2009), 'Do House Prices Reflect Fundamentals? Aggregate and Panel Data Evidence', *Journal of Housing Economics* **18**(2), 140–149.

Muellbauer, J. and Murphy, A. (1997), 'Booms and Busts in the U.K. Housing Market', *The Economic Journal* **107**(445), 1701–1727.

NOU (2002), Boligmarkedene og boligpolitikken. Norges Offentlige Utredninger, 2002:2.

NOU (2011), Bedre rustet mot finanskriser. Norges Offentlige Utredninger, 2011:1.

Pirounakis, N. G. (2013), *Real Estate Economics: A Point-to-Point Handbook*, Routledge, London.

Poterba, J. M. (1984), 'Tax Subsidies to Owner-Occupied Housing: An Asset-Market Approach', *The Quarterly Journal of Economics* **99**(4), 729–752.

Reinhart, C. M. and Rogoff, K. (2009), *This Time Is Different: Eight Centuries of Financial Folly*, Princeton University Press, Princeton, New Jersey.

Shiller, R. J. (2000), *Irrational Exuberance*, Princeton University Press, Princeton, New Jersey.

Sørensen, P. B. and Whitta-Jacobsen, H. J. (2010), *Introducing Advanced Macroeconomics: Growth and Business Cycles*, second edn, McGraw-Hill, London.

Stiglitz, J. E. (1990), 'Symposium on Bubbles', *Journal of Economic Perspectives* **4**(2), 13–18.

Tirole, J. (1988), *The Theory of Industrial Organization*, MIT Press, Cambridge, Massachusetts.

Tobin, J. (1969), 'A General Equilibrium Approach to Monetary Theory', *Journal of Money, Credit and Banking* **1**(1), 15–29.

Vale, B. (2004), The Norwegian Banking Crisis, *in* T. G. Moe, J. A. Solheim and B. Vale, eds, 'The Norwegian Banking Crisis', number 33 *in* 'Occasional Papers', Norges Bank, Oslo, chapter 1, pp. 1–21.

Appendix A Unit Root Tests for Stationarity

Appendix A investigates the stationarity properties of some of the key regional variables used in the housing market analysis.

Real household debt

The test results reported in Table 18 indicate that the series is an $I(2)$ process. Despite these formal test results, the debt is treated as an $I(1)$ variable in the econometric analysis.

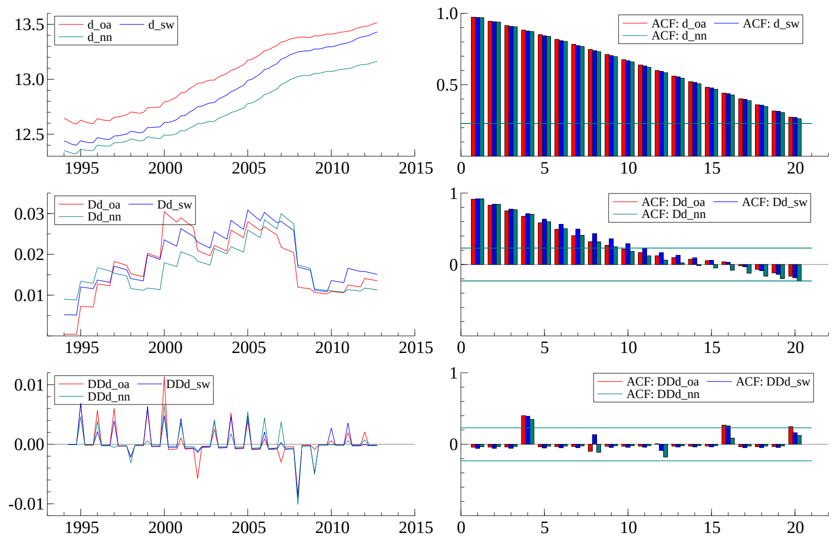


Figure 9: The stationarity properties of the real debt series in the three regions.

Table 18: Unit root ADF test for the regional debt series

Variabel	Lags	Constant	Trend	Seasonal	τ-ADF	5%-critical value
d^{OA}	2	✓	✓	–	−1.13	−3.47
d^{SW}	2	✓	✓	–	−1.65	−3.47
d^{NN}	2	✓	✓	–	−1.40	−3.47
Δd^{OA}	1	✓	–	–	−1.89	−2.90
Δd^{SW}	1	✓	–	–	−1.58	−2.90
Δd^{NN}	1	✓	–	–	−1.41	−2.90
$\Delta^2 d^{OA}$	1	✓	–	–	−6.29	−2.90
$\Delta^2 d^{SW}$	1	✓	–	–	−6.53	−2.90
$\Delta^2 d^{NN}$	2	✓	–	–	−5.11	−2.90
Estimation period:	1995Q1(2)[a](3)[b]– 2012Q4					

[a] For Δd^i the estimation period starts in 1995Q2.
[b] For $\Delta^2 d^i$ the estimation period starts in 1995Q3.

Real household income

Regional real household income is an $I(1)$ process.

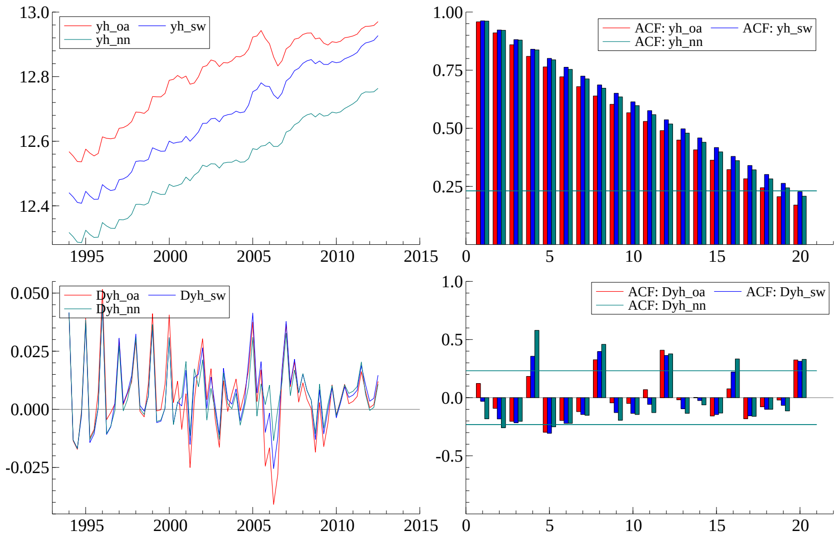


Figure 10: The stationarity properties of the real income series in the three regions.

Table 19: Unit root ADF test for the regional income series

Variabel	Lags	Constant	Trend	Seasonal	τ-ADF	5%-critical value
yh^{OA}	4	✓	✓	✓	-1.81	-3.47
yh^{SW}	2	✓	✓	✓	-3.03	-3.47
yh^{NN}	3	✓	✓	✓	-2.32	-3.47
Δyh^{OA}	3	✓	–	✓	-4.21	-2.90
Δyh^{SW}	2	✓	–	✓	-5.37	-2.90
Δyh^{NN}	2	✓	–	✓	-6.36	-2.90
Estimation period:	1995Q1–2012Q3					

Real wealth

The regional wealth series is an $I(1)$ process.

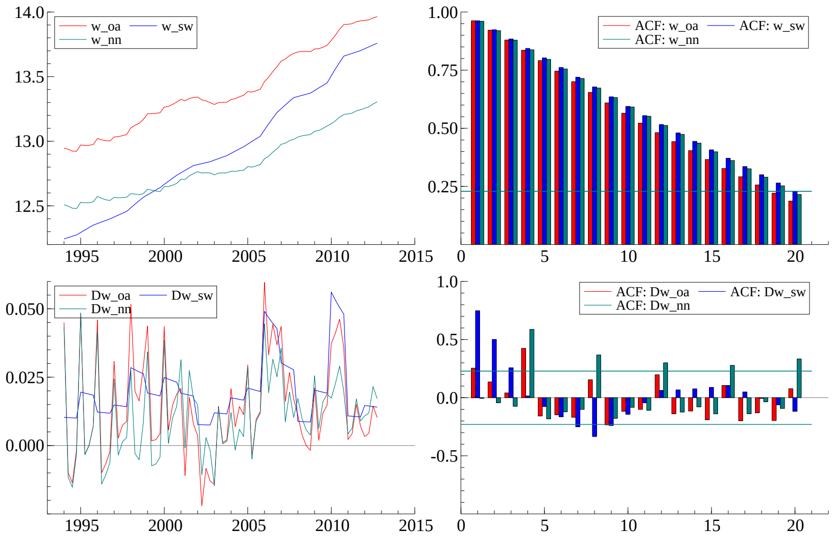


Figure 11: The stationarity properties of the real wealth series in the three regions.

Table 20: Unit root ADF test for the regional wealth series

Variabel	Lags	Constant	Trend	Seasonal	τ-ADF	5%-critical value
w^{OA}	3	✓	–	–	-0.03	-2.90
w^{SW}	2	✓	–	–	-0.27	-2.90
w^{NN}	2	✓	–	–	-1.68	-2.90
Δw^{OA}	2	✓	–	–	-3.68	-2.90
Δw^{SW}	4	✓	–	–	-3.05	-2.90
Δw^{NN}	2	✓	–	–	-4.58	-2.90
Estimation period:	1995Q1(2)[a]–2012Q3					

[a] For Δw^i the estimation period starts in 1995Q2.

Real interest rate

The aggregate real after-tax interest rate is a stationary $I(0)$ series.

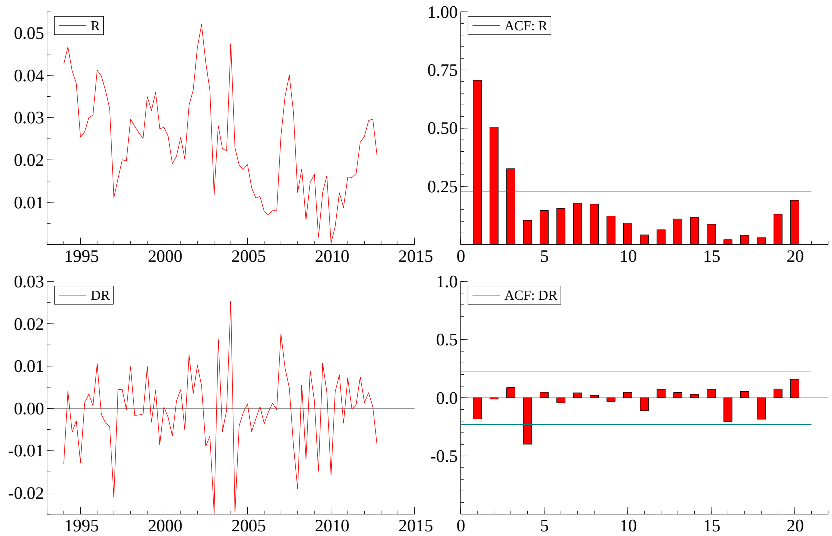


Figure 12: The stationarity properties of the aggregate real interest rate variable.

Table 21: Unit root ADF test for the aggregate real interest rate variable

Variabel	Lags	Constant	Trend	Seasonal	τ-ADF	5%-critical value
R	2	✓	–	–	-3.55	-2.90
ΔR	1	✓	–	–	-6.99	-2.90
Estimation period:	1994Q1–2012Q4					

Housing stock

The aggregate housing stock variable is an $I(2)$ series, however, it will be treated as an $I(1)$ series in the analysis.

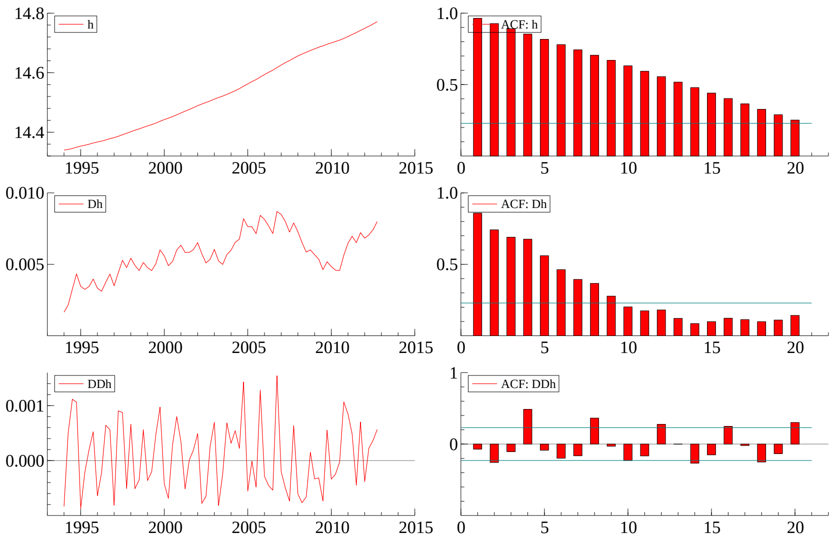


Figure 13: The stationarity properties of the aggregate housing stock variable.

Table 22: Unit root ADF test for the aggregate housing stock variable

Variabel	Lags	Constant	Trend	Seasonal	τ-ADF	5%-critical value
h	2	✓	✓	–	-2.25	-3.47
Δh	1	✓	–	–	-1.96	-2.90
$\Delta^2 h$	4	✓	–	–	-3.08	-2.90
Estimation period:	1994Q1–2012Q4					

Housing turnover

The aggregate housing turnover variable is an $I(1)$ series.

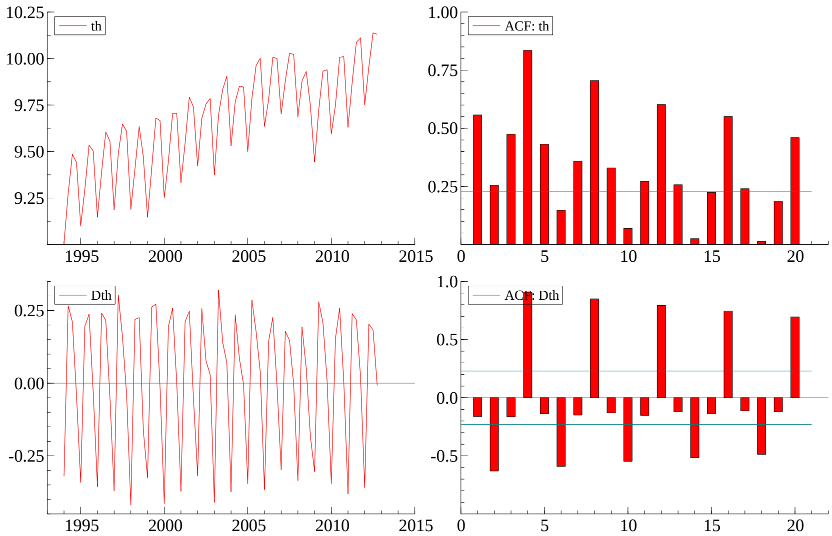


Figure 14: The stationarity properties of the aggregate housing turnover variable.

Table 23: Unit root ADF test for the aggregate housing turnover variable

Variabel	Lags	Constant	Trend	Seasonal	τ-ADF	5%-critical value
th	2	✓	–	✓	-1.39	-2.90
Δth	1	✓	–	✓	-6.66	-2.90
Estimation period:	1994Q1–2012Q4					

Appendix B Mis-specification Tests

As econometricians we want our econometric model to capture the data generating process (DGP), i.e. whatever mechanism in the real world that has produced the data. However, since one can never know the true DGP, a set of mis-specification tests can be conducted to see if the residuals in the model behaves significantly differently from what we would expect to see if the true residuals of the model were consistent with the classical assumptions of no autocorrelation, normality, and homoskedasticity. These classical assumptions is closely related to the Gauss-Markov theorem and the BLUE property of estimators. This appendix gives a brief overview of the standard mis-specification tests used in the econometric analysis of chapters 6 and 7, and explains which implications any violations of the classical conditions have on the estimation. A more thorough review is found in e.g. Doornik and Hendry (2009, chap. 18) and the references therein, and Bårdsen and Nymoen (2011, chap. 8) for a Norwegian exposition.

AR Test

The AR test is an LM test[37] for residual autocorrelation, which could be an issue when working with time-series data. In Table 2 it is reported as AR 1-5, indicating that it tests for autocorrelation up to order 5, i.e. five lags of the residuals. It is an F-test that tests the joint null hypothesis that $\hat{\varepsilon}_t$ is uncorrelated with $\hat{\varepsilon}_{t-j}$ for any j, against the alternative that $\hat{\varepsilon}_t$ is correlated with at least one $\hat{\varepsilon}_{t-j}$, where $j = 1, \ldots, 5$.

More formally, consider the econometric model

$$Y_t = \alpha_0 + \alpha_1 X_t + \varepsilon_t \tag{1}$$

Then, the AR test makes use of the auxiliary regression

$$\hat{\varepsilon}_t = \alpha_0 + \gamma_j \sum_{j=1}^{5} \hat{\varepsilon}_{t-j} + \alpha_1 X_t + \upsilon_t \tag{2}$$

where the null and alternative hypotheses are

$$H_0 : \gamma_1 = \gamma_2 = \gamma_3 = \gamma_4 = \gamma_5 = 0 \tag{3a}$$

$$H_A : \textit{At least one } \text{ of the } \gamma_j \text{ nonzero for} j = 1, \ldots, 5 \tag{3b}$$

[37]Lagrange multiplier test.

The null hypothesis is rejected if the test statistic is too high, and autocorrelation exists. The implications of autocorrelation depends on whether the regressors are exogenous or predetermined.[38][39] As Table 24 shows, inference based on t- and F-statistics when autocorrelation is present may result in misleading conclusions as the standard errors are wrong. Furthermore, in dynamic time-series models where the regressors are likely to be predetermined, autocorrelated residuals cause coefficient estimators to be both biased and inconsistent.

Table 24: Implications of autocorrelated disturbances

Autocorrelation		
X_t	$\hat{\alpha}_1$	$v\hat{a}r(\hat{\alpha}_1)$
Exogenous	unbiased, consistent	wrong
Predetermined	biased, inconsistent	wrong

Normality Test

The normality test, also known as the Jarque-Bera test, checks whether the null hypothesis of normally distributed residuals holds. It is based on two measures, skewness and kurtosis. Skewness refers to how symmetric the residuals are around zero, while kurtosis refers to the "peakedness" of the distribution. Failing to accept the null hypothesis means that the normality assumption does not hold, and inference based on t- and F-statistics is incorrect (because the corresponding distributions are not close to the normal distribution). Although the t- and F-distributions are no longer exact, they can still be a good approximation, and it becomes better as the sample size increases.

Hetero Test

The hetero test, or White test, is a test to check for heteroskedastic residuals, i.e. to verify the homoskedasticity assumption of the residuals having constant variance, $var(\varepsilon_i) = \sigma^2$. Formally, and with the basis in equation (1), the test involves the auxiliary regression

$$\hat{\varepsilon}_t = \alpha_0 + \alpha_1 X_t + \alpha_2 X_t^2 \tag{4}$$

If the model in equation (1) included more explanatory variables, then these variables and their squares would also be included in equation (4). The null and alternative hypotheses are

[38] A predetermined variable is a variable that is neither completely exogenous, not completely endogenous, as it is likely to be correlated with some of the residuals, and uncorrelated with others, i.e. $E(\varepsilon_t|\mathbf{X}_t) = 0$. See Bårdsen and Nymoen (2011, pg. 132)

[39] The standard notation of a dynamic model including a predetermined variable, would be to substitute X_t in equation (1) with Y_{t-1}.

$$H_0 : \alpha_1 = \alpha_2 = 0 \tag{5a}$$

$$H_A : \alpha_1 \neq 0 \text{ or } \alpha_2 \neq 0 \tag{5b}$$

The implications of heteroskedastic disturbances are summarized in Table 25. Since the errors are being incorrectly calculated when heteroskedasticity is present in the model, inference based on the usual test statistic may be misleading. Furthermore, in a dynamic model with predetermined regressors the coefficient estimators are biased, but consistent, i.e. the estimator produces an estimate which tends to its "true" value as the sample size increases, $\operatorname*{plim}_{n\to\infty} \hat{\alpha}_1 = \alpha_1$.

Table 25: Implications of heteroskedastic disturbances

Heteroskedasticity		
X_t	$\hat{\alpha}_1$	$v\hat{a}r(\hat{\alpha}_1)$
Exogenous	unbiased, consistent	wrong
Predetermined	biased, consistent	wrong

Appendix C Estimation Results From Anundsen and Jansen (2013*b*)

The trace test results given in Table 3 pg. 16 in Anundsen and Jansen (2013*b*).

Table 26: Trace test for cointegration over the sample 1986Q2–2008Q4

Eigenvalue: λ_i	H_0	H_A	λ_{trace}	5%-critical value
0.39	$r = 0$	$r \geq 1$	86.59	64.48
0.22	$r \leq 1$	$r \geq 2$	41.74	40.95
0.19	$r \leq 2$	$r \geq 3$	18.82	20.89
Diagnostics	Test statistic	Value (p-value)		
Vector AR 1-5 test:	F(45,146)	1.06 (0.39)		
Vector Normality test:	$\chi^2(6)$	7.78 (0.26)		
Vector Hetero test:	F(270,247)	1.03 (0.42)		
Estimation period:	1986Q2-2008Q4			

The estimation results given in Table 4 pg. 17 in Anundsen and Jansen (2013*b*) are given in Table 27. The justification for A&J's identification strategy and the imposed overidentifying restrictions are given in their paper.

Table 27: Testing steady-state hypotheses over the sample 1986Q2–2008Q4

The just identified housing price and debt equation are defined by

$$ph = \beta_{d,1}d + \beta_{yh,1}yh + \beta_{h,1}h + \beta_{R,1}R + \beta_{t,1}t$$

$$d = \beta_{ph,2}ph + \beta_{yh,2}yh + \beta_{R,2}R + \beta_{th,2}th + \beta_{h,2}h + \beta_{t,2}t$$

Panel 1: Testing no trend ($\beta_{t,1} = \beta_{t,2} = 0$)

$$ph = \underset{(0.07)}{0.76}d + \underset{(0.21)}{1.39}yh - \underset{(0.37)}{2.00}h - \underset{(0.85)}{0.13}R$$

$$d = 1.53ph - \underset{(0.17)}{1.45}yh - \underset{(1.40)}{0.71}R + \underset{(0.05)}{0.09}th + \underset{(0.07)}{1.53}h$$

$$LogL = 842.834,\ \chi^2(2) = 3.81(0.15)$$

Panel 2: No effect of real after tax interest rate on house prices ($\beta_{R,1} = 0$)

$$ph = \underset{(0.08)}{0.77}d + \underset{(0.22)}{1.43}yh - \underset{(0.40)}{2.07}h$$

$$d = 1.54ph - \underset{(0.18)}{1.48}yh - \underset{(0.40)}{0.54}R + \underset{(0.05)}{0.10}th + \underset{(0.07)}{1.54}h$$

$$LogL = 842.834,\ \chi^2(3) = 3.84(0.28)$$

Panel 3: No effect of disequilibrium housing prices in household debt

$$ph = \underset{(0.19)}{0.84}d + \underset{(0.65)}{1.67}yh - \underset{(1.18)}{2.58}h$$

$$d = 1.08ph - \underset{(0.85)}{1.18}yh - \underset{(2.35)}{3.98}R + \underset{(0.28)}{0.56}th + \underset{(0.30)}{1.08}h$$

$$LogL = 842.276,\ \chi^2(4) = 4.95(0.29)$$

Panel 4: No effect of real disposable income on household debt ($\beta_{yh,2} = 0$)

$$ph = \underset{(0.19)}{0.86}d + \underset{(0.64)}{1.42}yh - \underset{(1.16)}{2.33}h$$

$$d = 0.78ph - \underset{(1.87)}{2.83}R + \underset{(0.15)}{0.24}th + \underset{(0.15)}{0.78}h$$

$$LogL = 841.323,\ \chi^2(5) = 6.86(0.23)$$

Panel 5: Imposing weak exogeneity of income with respect to the long-run coefficients

$$ph = \underset{(0.19)}{0.98}d + \underset{(0.63)}{1.69}yh - \underset{(1.15)}{3.03}h$$

$$d = 0.76ph - \underset{(1.79)}{2.74}R + \underset{(0.15)}{0.28}th + \underset{(0.16)}{0.76}h$$

$$\alpha_{1,ph} = \underset{(0.04)}{-0.24},\ \alpha_{1,d} = \underset{(0.03)}{-0.10},\ \alpha_{2,d} = \underset{(0.01)}{-0.04}$$

$$LogL = 840.529,\ \chi^2(7) = 8.44(0.30)$$

The sample is 1986Q2 to 2008Q4, 91 observations